The Biometric Society

BIOMETRICS

FOUNDED BY THE BIOMETRICS SECTION OF THE AMERICAN STATISTICAL ASSOCIATION

TABLE OF CONTENTS

Number 3 September 1957 Volume 13

The Biometric Society is an international society devoted to the mathematical and statistical aspects of biology. Biologists, mathematicians, statisticians and others interested in its objectives are invited to become members. Through its regional organizations the Society sponsors regional and local meetings. National secretaries serve the interests of members in Denmark, India, Japan, the Netherlands, Sweden and Switzerland, and there are many members at large.

Rates (in U.S.A. currency) for full membership in the Society for 1957 including dues and a subscription to this journal are: for residents of Canada and the United States $7.00; for others $4.50. Members of the American Statistical Association who are currently subscribing to Biometrics through that organization may become members of the Biometric Society on payment of $3.00 annual dues if resident in the United States or Canada, and of $1.75 annual dues if resident elsewhere. Information concerning the Society can be obtained from its Secretary, M. J. R. Healy, Statistic Department, Rothamsted Experimental Station, Harpenden, Herts, England.

Annual subscription for non-members of the Biometric Society is $7.00, payable to the Managing Editor, *Biometrics*, Dept. of Statistics, Virginia Polytechnic Institute, Blacksburg, Virginia.

Second-class mailing privileges authorized at Knoxville, Tenn. Additional entry at Richmond, Va. Business Office: 509 West Hill Road, Knoxville 19, Tennessee. *Biometrics* is published quarterly—in March, June, September and December.

ANALYSIS OF COVARIANCE: ITS NATURE AND USES*

WILLIAM G. COCHRAN
The Johns Hopkins University, Baltimore, Maryland, U.S.A.

1. INTRODUCTION

This paper is intended as an introduction to the subsequent papers in this issue. It discusses the nature and principal uses of the analysis of covariance, and presents the standard methods and tests of significance.

As Fisher [1934] has expressed it, the analysis of covariance "combines the advantages and reconciles the requirements of the two very widely applicable procedures known as regression and analysis of variance." This dual role can be illustrated by a two-way classification in which rows represent treatments, and columns represent blocks or replications. The typical mathematical model appropriate to the analysis of covariance is

$$y_{ij} = \mu + \tau_i + \rho_j + \beta(x_{ij} - x_{..}) + e_{ij} \qquad (1)$$

Here y_{ij} is the yield or response, while x_{ij} is an auxiliary variate, sometimes called the *concomitant variate* or *covariate*, on which y_{ij} has a linear regression with regression coefficient β. The constants μ, τ_i and ρ_j are the true mean response and the effects of the ith treatment and jth replication, respectively. The residuals e_{ij} are random variates, assumed in standard theory to be normally and independently distributed with mean zero and common variance.**

From the viewpoint of analysis of variance, equation (1) may be rewritten as

$$y_{ij} - \beta(x_{ij} - x_{..}) = \mu + \tau_i + \rho_j + e_{ij} \qquad (2)$$

In this form, (2) is the typical equation for an analysis of variance of the quantities

$$y_{ij} - \beta(x_{ij} - x_{..})$$

*Paper No. 319, Department of Biostatistics.

**The symbols $x_{..}$, $y_{..}$ denote overall means, while $x_{i.}$, $y_{i.}$ denote treatment means.

These are the deviations of y_{ij} from its linear regression on x_{ij} , or the values of y_{ij} after adjustment for this linear regression. In this setting, τ_i may be regarded as the true effect of the ith treatment on y_{ij} , after adjustment for the linear regression on the covariate x_{ij} . Thus the technique enables us to remove that part of an observed treatment effect which can be attributed to a linear association with the x_{ij} .

When the objective of the analysis is to fit a regression of y on x, the parameters τ_i and ρ_j in equation (1) represent "nuisance" parameters, included in the mathematical specification in order to make it realistic. In this way the analysis of covariance extends the study of regression relationships to data of complex structure in which the nature of the regression is at first sight obscured by structural effects like the τ_i and ρ_j.

2. PRINCIPAL USES

These may be grouped under several headings.

2.1 *To increase precision in randomized experiments.* This is probably the most frequent application. The covariate x is a measurement, taken on each experimental unit before the treatments are applied, which is thought to predict to some degree the final response y on that unit. The first illustration of the covariance method in the literature was of this type (Fisher [1932]). The variate x was the yield of tea per plot in a period preceding the start of the experiment, while y was the tea yield at the end of a period of application of treatments (in this illustration, the treatments were "dummy"). Adjustment of the responses y for their regression on x removes the effects of variations in initial yields from the experimental errors, insofar as these effects are measured by the linear regression. In this example these effects might be due either to inherent differences in the tea bushes or to soil fertility differences that were permanent enough to persist during the course of the experiment.

With a linear regression equation, the gain in precision from the covariance adjustment depends primarily on the size of the correlation coefficient ρ between y and x on experimental units (plots) that receive the same treatment. If σ_y^2 is the experimental error variance when no covariance is employed, the adjustments reduce this variance to a value which is effectively about

$$\sigma_y^2(1 - \rho^2)\left\{1 + \frac{1}{f_e - 2}\right\}$$

where f_e is the number of error d.f. The factor involving f_e is needed to take account of errors in the estimated regression coefficient. If

ρ is less than 0.3 in absolute value, the reduction in variance is inconsequential, but as ρ mounts towards unity, sizeable increases in precision are obtained. In Fisher's example ρ was 0.928, reflecting a high degree of stability in relative yield of a plot from one period to another. The adjustment reduced the error variance roughly to a fraction $\{1 - (0.928)^2\}$, or about one-sixth, of its original value. Some of the most spectacular gains in precision from covariance have occurred in situations like this, in which the covariate represents an initial calibration of the responsiveness of the experimental units.

In this use the function of covariance is the same as that of local control (pairing and blocking). It removes the effects of an environmental source of variation that would otherwise inflate the experimental error. When the relation between y and x is linear, covariance and blocking are about equally effective.* If instead of using covariance we can group the units into replications such that the x values are equal within a replication, this blocking also reduces the error variance to $\sigma_y^2(1 - \rho^2)$.

The potentialities of preliminary measurements as a means of increasing precision have frequently been recognized by experimenters. In animal feeding experiments the response is taken as gain in weight (final-initial weight) rather than final weight itself. Insulin may be assayed from the drop in blood sugar (initial reading—reading 3 hours after injection of insulin) instead of from the 3 hour reading. The weight of a treated muscle on the right side of the body may be taken as a percentage of the weight of the corresponding untreated muscle on the left side. Such adjustments make the best use of the covariate only when the relation between y and x is exactly that implied by the adjustment. In the animal feeding and insulin examples, the assumption is that $\beta = 1$; in the muscles, that y/x is independent of x and has constant variance. If these assumptions do not hold, the adjustment falls short of the optimum and sometimes is worse than no adjustment at all. By a covariance analysis, the experimenter can utilize his knowledge or speculations about the general nature of the relation between y and x, but still leave flexibility in the process by including parameters like β that are estimated from the data. Incidentally, he can verify from the covariance analysis whether a specific simple adjustment like the use of $(y - x)$ is good enough, as it sometimes is.

In a covariance analysis, the preliminary variate x may be measured on a completely different scale from that of the response y,—a situation in which the experimenter would have difficulty in creating a "home-made" method of adjustment. Bartlett [1937] used a visual estimate

*See p. 281.

of the degree of saltiness of the soil to adjust cotton yields. Federer and Schlottefeldt [1954] used the serial order (1, 2, $\cdots$ 7) of the plot within a replication as a basis for a quadratic regression adjustment of tobacco data, thereby removing the effects of an unexpected gradient in fertility within the replications. Similarly, the reading performances of children under different methods of instruction may be adjusted for variations in their initial I.Q.'s. Note also that x need not be a direct causal agent of y—it may, for instance, merely reflect some characteristic of the environment that also influences y.

When covariance is used in this way, it is important to verify that the treatments have had no effect on x. This is obviously true when the x's were measured before the treatments were applied. But sometimes the x variates are measured after treatments have been applied, as when plant number shortly before harvest is used to adjust crop yields for uneven growth, or as happened in the index of saltiness used by Bartlett. When the treatments do affect the x-values to some extent, the covariance adjustments take on a different meaning. They no longer merely remove a component of experimental error. In addition, they distort the nature of the treatment effect that is being measured. If the higher yields given by superior treatments are due mostly to their effects in increasing numbers of plants, a covariance adjustment, which attempts to measure what yields would be if plant numbers were equal for all treatments, may remove most of the real treatment effect. The F-test of treatments against error for the x-variate is helpful when there is doubt whether treatments have had some effect on x.

2.2. *To remove the effects of disturbing variables in observational studies.* In fields of research in which randomized experiments are not feasible, we may observe two or more groups differing in some characteristic, in the hope of discovering whether there is an association between this characteristic and a response y. Examples are differences in heights of urban and rural school children, differences in illness rates between tenants of public and slum housing, and differences in expenditures for luxuries between clerical and manual workers. In observational studies it is widely realized that an observed association, even if statistically significant, may be due wholly or partly to other disturbing variables x_1 , x_2 $\cdots$ in which the groups differ. Where feasible, a common device, analogous to blocking in randomized experiments, is to match the groups for the disturbing variables thought to be most important. In the same way, a covariance adjustment may be tried for x-variables that have not been matched.

In a comparison of the heights of children from two different types of school, Greenberg [1953] found that the two groups differed slightly, though not significantly, in mean age. A covariance adjustment for age resulted in a more sensitive comparison of the heights. As a more complex example, Day and Fisher [1937] adjusted the log S. D. of leaf length (used as a measure of within-species variability) for fluctuations in length, breadth and thickness of leaves in comparing populations of *Plantago maritima* from different regions.

In observational studies covariance can perform two distinct functions. One is to remove bias. To illustrate, it follows from model (1) that the *unadjusted* difference $(y_{i.} - y_{j.})$ between the means of two groups is

$$y_{i.} - y_{j.} = \tau_i - \tau_j + \beta(x_{i.} - x_{j.}) + e_{i.} - e_{j.}$$

If the two groups have not been matched for x, the difference $(x_{i.} - x_{j.})$ may reflect a real difference in their x-distributions, being much larger than can be accounted for by within-group variations. The term $\beta(x_{i.} - x_{j.})$ is then of the nature of a bias which if allowed to remain will render tests of significance and confidence limits invalid. If model (1) applies to the data at hand, a covariance adjustment removes their bias. Most users of covariance in observational studies would, I think, regard coping with bias as its primary function. However, even if there are no real differences between the x-distributions in the two groups, so that there is no danger of bias, covariance may still be used to increase the precision of the comparison as in the applications in section 2.1.

Unfortunately, observational studies are subject to difficulties of interpretation from which randomized experiments are free. Although matching and covariance have been skillfully applied, we can never be sure that bias may not be present from some disturbing variable that was overlooked. In randomized experiments, the effects of this variable are distributed among the groups by the randomization in a way that is taken into account in the standard tests of significance. There is no such safeguard in the absence of randomization.

Secondly, when the x-variables show real differences among groups—the case in which adjustment is needed most—covariance adjustments involve a greater or less degree of extrapolation. To illustrate by an extreme case, suppose that we were adjusting for differences in parents' income in a comparison of private and public school children, and that the private-school incomes ranged from \$10,000–\$12,000, while the public-school incomes ranged from \$4,000–\$6,000. The covariance would adjust results so that they allegedly applied to a mean income

of \$8,000 in each group, although neither group has any observations in which incomes are at or even near this level.

Two consequences of this extrapolation should be noted. Unless the regression equation holds in the region in which observations are lacking, covariance will not remove all the bias, and in practice may remove only a small part of it. Secondly, even if the regression is valid in the no man's land, the standard errors of the adjusted means become large, because the standard error formula in a covariance analysis takes account of the fact that extrapolation is being employed (although it does not allow for errors in the form of the regression equation). Consequently the adjusted differences may become insignificant statistically merely because the adjusted comparisons are of low precision.

When the groups differ widely in x, these difficulties imply that the interpretation of an adjusted analysis is speculative rather than soundly based. While there is no sure way out of this difficulty, two precautions are worth observing.

(i) Consider what internal or external evidence exists to indicate whether the regression is valid in the region of extrapolation. Sometimes the fitting of a more complex regression formula serves as a partial check.

(ii) Examine the standard errors of the adjusted group means, particularly when differences become non-significant after adjustment. Confidence limits for the difference in adjusted means will reveal how precise or imprecise the adjusted comparison is.

2.3 *To throw light on the nature of treatment effects.* This application is closely related to the previous one. In a randomized experiment, the effects of several soil fumigants on eelworms, which attack some English farm crops, were compared. After the treatments had been given time to exert their effects, the numbers of eelworm cysts per plot and the yields of the crop, spring oats, were both recorded. Significant effects were produced on both eelworms and oats. It would be of interest to discover whether the reductions in numbers of eelworms were the causal mechanism in producing the observed differences in oats yields. If the treatment effects on the oats (y) disappear after adjusting by covariance for differences in the numbers of eelworms (x), this suggests, at least at first sight, that the treatment differences are simply a reflection of the differences produced by the fumigants on eelworm numbers. There are numerous instances of this kind in which a concomitant variable might be in part the agent through which the treatments produce their effects on the principal response. A covariance analysis offers the possibility of exploring whether this is so.

Here again, however, there are difficulties which restrict the utility of this ingenious tool. These are analyzed by Fairfield Smith in a later paper in this issue. Since in most of these applications the treatments will have produced significant effects on x, there is the problem of extrapolation discussed in the previous section. In addition, as Fairfield Smith illustrates, the interpretation of the adjusted y averages requires careful study. Sometimes these averages have no physical or biological meaning of interest to the investigator, and sometimes they do not have the meaning that is ascribed to them at first glance.

2.4 *To fit regressions in multiple classifications.* The simplest situation, discussed in elementary text books, involves a single classification. By standard techniques we can (i) fit a separate regression of y on x within each class, (ii) test whether the slopes or positions of the lines differ from class to class, and (iii) if advisable, make a combined estimate of a common slope. As an example of a regression in a row $\times$ column classification, the regression of wheat yield on shoot height, number of plants and number of ears was worked out from a series of growth studies on wheat conducted in Britain (Cochran [1938]). Each quartet of observations (y, x_1 , x_2 , x_3) represented the mean over several plots at one station in one year. In order that soil, geographic and seasonal factors would be adequately sampled, data were obtained from 7 stations for each of 5 years, making 35 quartets. Consequently it was necessary to fit constants for the mean yield of each station and the mean yield of each year.

2.5 *To analyze data when some observations are missing.* An interesting by-product of the covariance method, first pointed out by Bartlett (1937), is that it may be used to compute the exact analysis of variance when some observations are missing. To each missing observation we assign any convenient value (*e.g.* 0, 5, or 100) and introduce a dummy x-variate that takes the value[1] for the missing unit, and 0 for all other units. The standard covariance computations then give the correct least squares estimates of the treatment means and the exact F- and t-tests. This method is probably slower than the insertion of a missing value by the Yates formula [1933], but it is useful (i) with unfamiliar classifications where the Yates formula has not been worked out and (ii) where exact F- and t-tests are important, since the Yates' method gives only approximate tests.

Covariance can also be used (Nair [1939]) to estimate individual yields of a group of plots whose produce has inadvertently been combined, so that only a total over the group is known.

3. THE STANDARD COMPUTATIONS

The computations, which will be reviewed briefly, are essentially the same for all mathematical models in which a single regression and a single residual variance are postulated. These cases include the simpler multiple classifications and experimental designs (randomized blocks, cross-over designs, latin squares) as well as balanced and partially balanced incomplete block designs without recovery of inter-block information. Separate discussion is required for hierarchical classifications involving more than one regression equation or more than one residual variance in the specification, as with split-plot designs or incomplete blocks when inter-block information is recovered.

The backbone of the standard procedure is an analysis of sums of squares and products into the Treatments and Error components. Table 1 shows the notation employed

TABLE 1

SUMS OF SQUARES AND PRODUCTS

	D. f.	(x^2)	(xy)	(y^2)
Treatments	$(t-1)$	T_{xx}	T_{xy}	T_{yy}
Error	f_e	E_{xx}	E_{xy}	E_{yy}
Sum	$t-1+f_e$	S_{xx}	S_{xy}	S_{yy}

A line is added giving the sums for treatments and error. Thus $S_{xx} = T_{xx} + E_{xx}$, etc.

The error s.s. for y is now divided into two parts: the s.s. for regression on x (1 d.f.) and the s.s. of deviations (Table 2). The same subdivision is made for the Sums.

TABLE 2

PARTITION OF E_{xx} AND S_{xx} INTO COMPONENTS FOR REGRESSION AND FOR DEVIATIONS

	(y^2)	Regression		Deviations		
		D. f.	S. s.	D. f.	S. s.	M. s.
Error	E_{xx}	1	E_{xy}^2/E_{xx}	$f_e - 1$	$E_{yy} - E_{xy}^2/E_{xx}$	s_e^2
Sum	S_{xx}	1	S_{xy}^2/S_{xx}	$t + f_e - 2$	$S_{yy} - S_{xy}^2/S_{xx}$	
Treatments (by subtraction)				$t-1$	$T_{yy} - S_{xy}^2/S_{xx} + E_{xy}^2/E_{xx}$	s_t^2

The reduced s.s. for Treatments is obtained by subtracting the deviations s.s. for Error from that for the Sum (last line of Table 2).

The items of information most commonly wanted from this analysis are obtained as follows.

(i) The regression coefficient β. This is estimated from the Error line; $b = E_{xy}/E_{xx}$. The estimated standard error of b is $s_e/\sqrt{E_{xx}}$, with $(f_e - 1)$ d.f.

(ii) The adjusted estimate of a treatment effect. In the simplest experimental designs, the *unadjusted* estimate for treatment i is simply the mean $y_{i.}$ of all observations having this treatment. The adjusted estimate is

$$y'_{i.} = y_{i.} - b(x_{i.} - x_{..}) \tag{1}$$

(iii) The estimated standard error of any linear function $L = \sum g_i y'_{i.}$ of the adjusted treatment means is

$$s_L = s_e \sqrt{\frac{\sum g_i^2}{r} + \frac{[\sum g_i(x_{i.} - x_{..})]^2}{E_{xx}}} \tag{2}$$

with $(f_e - 1)$ d.f., where r is the number of replications. In particular, the standard error of the difference between two adjusted treatment means is, putting $g_i = 1$, $g_j = -1$, and all other g's $= 0$,

$$s.e.(y'_{i.} - y'_{j.}) = s_e \sqrt{\frac{2}{r} + \frac{(x_{i.} - x_{j.})^2}{E_{xx}}} \tag{3}$$

(iv) For a test of the null hypothesis that all treatment effects are equal, we compute $F = s_t^2/s_e^2$, where s_t^2 and s_e^2 are the mean squares found in Table 2. This ratio has $(t - 1)$ and $(f_e - 1)$ d.f.

4. NATURE OF THE COVARIANCE ADJUSTMENT

The structure of the covariance adjustment, $-b(x_{i.} - x_{..})$, is in accord with common sense: $(x_{i.} - x_{..})$ measures the amount by which the x-value for this treatment exceeds the average x-value, while b measures the change in y expected to accompany unit change in x. In a specific application, the sizes and directions of the adjustments are determined by the data. In this respect a covariance adjustment differs markedly from the type of arbitrary adjustment that has sometimes earned a dubious reputation for the whole process of adjustment. It is not true, however, that a covariance adjustment is entirely objective, since the investigator must choose the type of regression equation (e.g. linear, quadratic, linear in log x) from which the adjustment is derived. Moreover, the x-variables do not always measure what we would like to think they measure. One occasionally meets extravagant

claims for a covariance adjustment, as for example that data have been "adjusted to equalize socio-economic status," when what has actually been done is to adjust for a linear regression on a crude social rating of the father's occupation as reported by anyone who happens to be at home when the interviewer calls.

From equation (3), the estimated variance of the difference between two adjusted means, when averaged over all pairs of means, works out as

$$\frac{2s_e^2}{r}\left\{1 + \frac{t_{xx}}{E_{xx}}\right\} \tag{4}$$

where $t_{xx} = T_{xx}/(t - 1)$ is the treatments *mean square* for x. If t-tests are to be made between several pairs of means, this expression may be used, as an approximation, for the variance of the difference between any pair (Finney [1946]). This saves the labor of computing a separate standard error for each pair, as is required by the exact formula (3). This device is not recommended when the treatments produce significant effects on x, because the variance of the difference may be substantially greater for some pairs than for others, so that the use of a single average variance becomes unsatisfactory.

More generally, the quantity

$$s_e^2\left\{1 + \frac{t_{xx}}{E_{xx}}\right\} \tag{5}$$

may be regarded as the effective error variance *per unit* in a covariance analysis, where the term in brackets is an allowance for sampling errors in b. In a completed experiment, the gain in precision from covariance can be estimated by comparing (5) with $s_y^2 = E_{yy}/f_e$, the error mean square for y in its analysis of variance in Table 1. This comparison ignores the loss of 1 d.f. from the error which occurs with a covariance adjustment. The effect of this loss on the sensitivity of the t-tests is small even with only 5 d.f. in error.

5. THEORY OF THE TECHNIQUE

The theory for the simplest designs will be illustrated by the row $\times$ column classification, with treatments as rows and replications as columns. The model is

$$y_{ij} = \mu + \tau_i + \rho_j + \beta(x_{ij} - x_{..}) + e_{ij} \tag{6}$$

Following the method of least squares, the unknown parameters are estimated by minimizing

$$\sum_{i,j} \{y_{ij} - m - t_i - r_j - b(x_{ij} - x_{..})\}^2 \tag{7}$$

Since t_i need measure only the difference between the effect of the ith treatment and the general mean, we may assume

$$\sum_i t_i = 0; \qquad \sum_j r_j = 0.$$

The estimates. The algebra involved in finding the estimates can be reduced by introducing the variable

$$x'_{ij} = x_{ij} - x_{i.} - x_{.j} + x_{..} \tag{8}$$

Those familiar with the analysis of variance will recognize x'_{ij} as the contribution of the (i, j)th observation to the error of x in its analysis of variance. The properties of x'_{ij} that will be useful are

$$\sum_i x'_{ij} = 0; \qquad \sum_j x'_{ij} = 0 \tag{9}$$

$$\sum_{i,j} x'^2_{ij} = E_{xx}\,; \qquad \sum_{i,j} x'_{ij} y_{ij} = E_{xy} \tag{10}$$

The second relation in (10) is less familiar than the others, but is easily verified.

Now the given prediction equation

$$y_{ij} = m + t_i + r_j + b(x_{ij} - x_{..})$$

becomes identical with the prediction equation

$$y_{ij} = m' + t'_i + r'_j + bx'_{ij}$$

if the new estimates m', t'_i , r'_j satisfy the relations

$$m = m'; \qquad t_i = t'_i - b(x_{i.} - x_{..}); \qquad r_j = r'_j - b(x_{.j} - x_{..}) \tag{11}$$

This may be verified by substitution. Further, since $\sum t_i = \sum r_j = 0$, it follows that $\sum t'_i = \sum r'_j = 0$.

Hence, instead of finding m, t_i , r_j and b so as to minimize (7), we can find m' , t'_i , r'_j and b to minimize

$$\sum_{i,j} (y_{ij} - m' - t'_i - r'_j - bx'_{ij})^2 \tag{12}$$

On expansion, (12) becomes

$$\sum_{i,j} (y_{ij} - m' - t'_i - r'_j)^2 - 2b \sum_{i,j} x'_{ij} y_{ij} + b^2 \sum_{i,j} x'^2_{ij}$$

Note that the omitted terms

$$\sum m' x'_{ij} , \qquad \sum t'_i x'_{ij} , \qquad \sum r'_j x'_{ij}$$

vanish because of relations (9).

Using (10), the quantity to be minimized is

$$\sum_{i,j} (y_{ij} - m' - t'_i - r'_j)^2 - 2bE_{xy} + b^2E_{xx} \tag{13}$$

The advantage of this result is that b is disentangled from the other unknowns. Differentiation with respect to b gives

$$b = E_{xy}/E_{xx} \tag{14}$$

The other unknowns, m', t'_i , r'_j , must be chosen so as to minimize the first term in (13). But this is exactly the minimization involved in an ordinary analysis of variance of y without covariance. Hence, by the standard results for the analysis of variance,

$$m' = y_{..} \; ; \qquad t'_i = y_{i.} - y_{..} \; ; \qquad r'_j = y_{.j} - y_{..}$$

Finally, from (11), the least squares estimates for the covariance analysis are

$$\left.\begin{aligned} m &= y_{..} \\ t_i &= y_{i.} - y_{..} - b(x_{i.} - x_{..}) \\ r_j &= y_{.j} - y_{..} - b(x_{.j} - x_{..}) \end{aligned}\right\} \tag{15}$$

Since the t_i represent deviations from the overall mean, the estimate used in practice is

$$m + t_i = y_{i.} - b(x_{i.} - x_{..}) = y'_{i.} \text{ , say.} \tag{16}$$

Standard errors. From (13), the residual sum of squares may be written

$$E_{yy} - 2bE_{xy} + b^2E_{xx} = E_{yy} - \frac{2E_{xy}^2}{E_{xx}} + \frac{E_{xy}^2}{E_{xx}} = E_{yy} - \frac{E_{xy}^2}{E_{xx}} \tag{17}$$

The d.f. are $(f_e - 1)$, since 1 d.f. is subtracted for the regression. Hence s_e^2 , the residual mean square, is computed as in Table 2.

To find the standard error of b, we may write

$$b = \frac{\sum x'_{ij}y_{ij}}{E_{xx}} = \frac{\sum x'_{ij}\{\mu + \tau_i + \rho_j + \beta(x_{ij} - x_{..}) + e_{ij}\}}{E_{xx}}$$

From the properties of the x'_{ij} , this reduces to

$$b = \beta + \frac{\sum x'_{ij}e_{ij}}{E_{xx}}$$

This equation expresses $(b - \beta)$ as a linear function of the random residuals e_{ij} . Hence

$$\sigma_b^2 = E(b - \beta)^2 = \frac{\sigma_e^2 \sum x_{ij}'^2}{E_{xx}^2} = \frac{\sigma_e^2}{E_{xx}}$$

In the same way we find, for an adjusted treatment mean,

$$y'_{i.} = m + t_i = \mu + \tau_i + e_{i.} - (x_{i.} - x_{..}) \frac{\sum x'_{ij} e_{ij}}{E_{xx}}$$

For a linear comparison between the adjusted means, it follows that

$$L = \sum g_i y'_{i.} = \sum g_i(\mu + \tau_i) + \sum g_i e_{i.} - \{\sum g_i(x_i - x_{..})\} \frac{\sum x'_{ij} e_{ij}}{E_{xx}}$$

By the properties of the x'_{ij}, the variate $\sum x'_{ij} e_{ij}$ is uncorrelated with any of the means $e_{i.}$. This gives

$$\sigma_L^2 = \sigma_e^2 \left\{ \frac{\sum g_i^2}{r} + \frac{[\sum g_i(x_{i.} - x_{..})]^2}{E_{xx}} \right\}$$

in agreement with the result in equation (2).

Note: The variate x'_{ij} can be used in the same way with all other designs (e.g. latin squares, incomplete blocks) to which the standard covariance theory applies. With an incomplete block design, for example, we define x'_{ij} as

$$x'_{ij} = x_{ij} - m_x - \hat{t}_{ix} - \hat{r}_{jx}$$

where $\hat{t}_{ix}$, $\hat{r}_{jx}$ are the estimates of the treatment and block effects, respectively, in the incomplete block analysis of variance of x. The rest of the algebraic development goes through without change.

Tests of significance. For the F-test, we quote the general theorem on the F-test in regression analysis, as applied to this problem. Let

$$D = \sum \{y_{ij} - m - t_i - r_j - b(x_{ij} - x_{..})\}^2$$

$$D'' = \sum \{y_{ij} - m'' - r''_j - b''(x_{ij} - x_{..})\}^2$$

where the constants are chosen in each case so as to minimize the corresponding sum of squares of deviations. Then if all τ_i are zero, the quantity

$$\frac{(D'' - D)}{(t - 1)} \Big/ \frac{D}{(f_e - 1)}$$

is distributed as F with $(t - 1)$ and $(f_e - 1)$ d.f. For proofs see, e.g., Yates [1938], Anderson and Bancroft [1952].

From (17) the denominator $D/(f_e - 1)$ has been shown to be s_e^2. By the same approach, D'' is the sum of squares of deviations from the "error" regression when only replication effects are eliminated in the analysis of variance. But in that event the "error" will be equal to

"treatments + error" in the present analysis. Hence the numerator of F, $(D'' - D)/(t - 1)$, is equal to s_t^2 as defined in Table 2.

6. THE REDUCED SUM OF SQUARES FOR TREATMENTS

The non-mathematical user of the analysis of covariance often finds it easy to rationalize the method by which b is computed and the formula for making the adjustments. The roundabout method by which the numerator of F is computed is, however, apt to appear mysterious. In this section the sum of squares $D'' - D$ in the numerator of F is examined in more detail. Following Fisher, this quantity will be called the *reduced* sum of squares for treatments, T_{yyR} .

Intuitively, one might expect this quantity to be a squared comparison among the adjusted treatment means

$$y'_{i.} = y_{i.} - b(x_{i.} - x_{..})$$

As is well-known, T_{yyR} is not the sum of squares of deviations of these adjusted treatment means. The latter sum of squares, T_{yyA} say, is

$$\begin{aligned} T_{yyA} &= r \sum \{(y_{i.} - y_{..}) - b(x_{i.} - x_{..})\}^2 \\ &= T_{yy} - 2bT_{xy} + b^2T_{xx} \qquad (18) \\ &= T_{yy} - 2bb_tT_{xx} + b^2T_{xx} \qquad (19) \end{aligned}$$

where $b_t = T_{xy}/T_{xx}$ is the regression coefficient as computed from the Treatments line in the analysis of variance in Table 1. From Table 2, on the other hand, the sum of squares T_{yyR} in the numerator of F is

$$\begin{aligned} T_{yyR} &= T_{yy} - \frac{(T_{xy} + E_{xy})^2}{T_{xx} + E_{xx}} + \frac{E_{xy}^2}{E_{xx}} \qquad (20) \\ &= T_{yy} - \frac{(b_tT_{xx} + bE_{xx})^2}{T_{xx} + E_{xx}} + b^2E_{xx} \qquad (21) \end{aligned}$$

Subtracting (19) from (21) and taking the common denominator $(T_{xx} + E_{xx})$, we find

$$T_{yyR} = T_{yyA} - \frac{T_{xx}^2(b_t - b)^2}{T_{xx} + E_{xx}} \qquad (22)$$

This result was first given by Yates [1934]. Equation (22) remains valid if T_{yyR} and T_{yyA} are the reduced and adjusted sums of squares for any specific component of the treatments sum of squares, provided that b_t and T_{xx} refer to this component. As Yates showed, the result provides an approximate short-cut method of making F-tests when we intend to test several components. The exact test involves going through the procedure in Table 2 separately for each component. We

can, however, easily calculate T_{yyA} for each component by relation (18). The F-values computed from T_{yyA} will all be too large, but the approximation is usually close if the treatments do not affect x, and only a few borderline F's need be recomputed by the exact method. Table 3 compares the approximate and exact F's in an experiment in which the treatments sum of squares was divided into 4 components, (data from Cochran and Cox [1957], s.s. rounded to 100's).

TABLE 3

THE APPROXIMATE F-TEST FROM THE ANALYSIS OF $(y - bx)^2$

	D.f.	Sums of squares x^2	xy	y^2	$(y - bx)^2$	F_A	F	$F_{.05}$
Treatments	8	292	− 92	1574	2571	4.51	4.16	2.22
Ave. L	1	31	−133	572	1062	14.90	14.53	4.12
Ave. Q	1	22	83	311	106	1.49	1.46	4.12
Dev. L	3	230	− 68	434	1205	5.63	5.05	2.87
Dev. Q	3	9	26	257	198	0.93	0.93	2.87
Error	36	1214	1893	5447	2495			

The columns on the right show the approximate ratio F_A, the exact F as found from the method in Table 2, and the tabular 5% value. Since none of the F_A values is near the 5% level, no recomputation would be necessary in practice in this example. (The column of mean squares of $(y - bx)^2$ has been omitted.)

The presence of sampling errors in b explains why T_{yyA} is not the correct sum of squares for the numerator of F. If the true β could be used in T_{yyA}, this sum of squares would give an exact F-test. The sampling error of b complicates the issue in two ways. Each adjusted treatment mean has a different variance, since

$$V(y'_{i.}) = \sigma_e^2\left\{\frac{1}{r} + \frac{(x_{i.} - x_{..})^2}{E_{xx}}\right\} \tag{23}$$

Further, the adjusted means for the ith and jth treatments are not independent, since b enters into both. Their covariance is

$$\text{cov}\,(y'_{i.}y'_{j.}) = \frac{\sigma_e^2(x_{i.} - x_{..})(x_{j.} - x_{..})}{E_{xx}} \tag{24}$$

Consequently, as would be expected from theory, it turns out that the correct numerator T_{yyR} is a quadratic form in the adjusted means,

$$T_{yyR} = \sum_{i,j} a_{ij}(y'_{i.} - y_{..})(y'_{j.} - y_{..})$$

where the matrix a_{ij} is the inverse of the variance-covariance matrix of the y_{ij} (apart from the factor σ_e^2). To find the a_{ij}, we may write, from (22),

$$T_{yyR} = r \sum (y'_{i.} - y_{..})^2 - \frac{T_{xx}^2(b_t - b)^2}{T_{xx} + E_{xx}} \tag{25}$$

Now

$$\begin{aligned} T_{xx}(b_t - b) &= T_{xy} - bT_{xx} \\ &= r \sum_i (x_{i.} - x_{..})\{(y_{i.} - y_{..}) - b(x_{i.} - x_{..})\} \\ &= r \sum_i (x_{i.} - x_{..})(y'_{i.} - y_{..}) \end{aligned}$$

Substituting this expression in (25) we find that

$$T_{yyR} = r \sum (y'_{i.} - y_{..})^2 - \frac{r^2[\sum (x_{i.} - x_{..})(y'_{i.} - y_{..})]^2}{T_{xx} + E_{xx}}$$

It follows that

$$\begin{aligned} a_{ii} &= r\left\{1 - \frac{r(x_{i.} - x_{..})^2}{T_{xx} + E_{xx}}\right\} \\ a_{ij} &= -\frac{r^2(x_{i.} - x_{..})(x_{j.} - x_{..})}{T_{xx} + E_{xx}} \qquad i \neq j \end{aligned} \tag{26}$$

The inverse of the matrix (a_{ij}) can be shown to be the covariance matrix of the y'_i as given by equations (23) and (24).

It is also instructive to compare T_{yyA} and T_{yyR} with the sum of squares of deviations of the treatment means $y_{i.}$ from the *treatments* regression on x, i.e. with the treatments sum of squares of $(y - b_t x)$. This sum of squares is

$$T_{yy.x} = T_{yy} - \frac{T_{xy}^2}{T_{xx}} = T_{yy} - b_t^2 T_{xx} \tag{27}$$

and has $(t - 2)$ d.f. From (19),

$$\begin{aligned} T_{yyA} &= T_{yy} - 2bb_t T_{xx} + b^2 T_{xx} \\ &= T_{yy.x} + (b_t - b)^2 T_{xx} \end{aligned} \tag{28}$$

using (27). Finally from (22),

$$\begin{aligned} T_{yyR} &= T_{yyA} - \frac{(b_t - b)^2 T_{xx}^2}{T_{xx} + E_{xx}} \\ &= T_{yy.x} + \frac{(b_t - b)^2 T_{xx} E_{xx}}{T_{xx} + E_{xx}} \end{aligned} \tag{29}$$

Equation (29) shows that the reduced sum of squares for treatments separates into two parts: (i) the sum of squares $T_{yy.x}$ with $(t - 2)$ d.f. which represents the deviations of the treatment means from their own regression, (ii) a single d.f. which compares the regression coefficients from the treatments and error lines in the analysis of variance. It may be shown that if there are no treatment effects, this sum of squares is distributed as $\chi^2\,\sigma_e^2$ with 1 d.f. In this way, equation (29) can be used to prove directly that, on the null hypothesis, T_{yyR} is distributed as $\chi^2\,\sigma_e^2$ with $(t - 1)$ d.f. In the same vein, equation (28) shows that T_{yyA} consists of $(t - 2)$ d.f. that are distributed as $\chi^2\,\sigma_e^2$. The remaining 1 d.f. is inflated by the factor $(T_{xx} + E_{xx})/E_{xx}$, as is seen by comparing (28) with (29).

To sum up, the sum of squares T_{yyR} in the numerator of the F-test is a quadratic form in the deviations $(y'_{i.} - y_{..})$ of the adjusted treatment means. The coefficients a_{ij} in this quadratic form are the inverse of the covariance matrix of the quantities $y'_{i.}$. As would be expected, T_{yyR} vanishes if all adjusted means are equal. Because of the presence of sampling errors in b, T_{yyR} does not equal the usual sum of squares of deviations T_{yyA} of the adjusted treatment means. T_{yyA} always gives too large a sum of squares, but if treatments do not affect x it is a useful approximation to T_{yyR} when several components of the treatments sum of squares are to be tested.

7. ASSUMPTIONS REQUIRED FOR THE ANALYSIS OF COVARIANCE

The assumptions required for valid use of the analysis of covariance are the natural extension of those for an analysis of variance, namely,

(i) Treatment, block and regression effects must be additive as postulated by the model,

(ii) The residuals e_{ij} must be normally and independently distributed with zero means and the same variance.

Although the effects of failures in these assumptions on the analysis of covariance as such do not appear to have been investigated, much of the related work on the analysis of variance carries over—for instance, that on the effects of non-normality or inhomogeneity of variance in the e_{ij}. The general precautions that have been given about the practical use of the analysis of variance should equally be observed in an analysis of covariance, see e.g. Cochran [1947].

Two assumptions that particularly involve the regression term in covariance should be noted. The treatment and the regression effects may not be additive. Bartlett [1937] pointed out that this danger might be present in the cotton-salt example already cited. On plots with a high salt content, the crop might be unable to respond to superior

fertilizers. Thus, in an extreme case, the treatment effects may be zero if x lies above a certain value. If this happens, the covariance adjustment may still improve the precision, but (i) the meaning of the adjusted treatment effects become cloudy, and (ii) if covariance is applied in a routine way, the investigator fails to discover the differential nature of the treatment effects—a point that might be important for practical applications.

Secondly, the covariance procedure assumes that the correct form of regression equation has been fitted. Perhaps the most common error to be anticipated is that linear regressions will be used when the true regression is curvilinear. In a randomized experiment in which treatments do not affect x, the randomization ensures that the usual interpretations of standard errors and tests of significance are not seriously vitiated, although fitting the correct form of regression would presumably give a larger increase in precision. The danger of misleading results is greater when x shows real differences from treatment to treatment. Later investigations by Fairfield Smith in this issue suggest, however, that this disturbance is serious only in rather extreme situations.

8. MULTIPLE COVARIANCE

No new difficulty is presented when there is more than one covariate, although the computations become more lengthy. Only the basic formulae will be given. With two covariates, x and z, the regression coefficients are obtained from the equations

$$E_{xx}b_{y.x} + E_{xz}b_{y.z} = E_{xy}$$

$$E_{zx}b_{y.x} + E_{zz}b_{y.z} = E_{zy}$$

If t-tests of the adjusted treatment means are wanted, it is advisable to compute the inverse of the E_{xz} matrix, say (c_{xz}).

In the simple orthogonal designs like randomized blocks, the adjusted mean of the ith treatment is estimated by

$$y'_{i.} = y_{i.} - b_{y.x}(x_{i.} - x_{..}) - b_{y.z}(z_{i.} - z_{..})$$

The variance of the difference between the ith and jth treatment effects is

$$s_e^2\left\{\frac{2}{r} + (x_{i.} - x_{j.})^2c_{xx} + 2(x_{i.} - x_{j.})(z_{i.} - z_{j.})c_{xz} + (z_{i.} - z_{j.})^2c_{zz}\right\}$$

The average variance over all pairs of treatments is

$$\bar{s}^2_{\text{diff}} = \frac{2}{r}s_e^2(1 + t_{xx}c_{xx} + 2t_{xz}c_{xz} + t_{zz}c_{zz})$$

where l_{xx} is the treatments mean square for x, etc. This is the extension of equation (4). If treatments do not affect x or z, the square root, $\bar{s}_{d(t)}$, may be used as an approximate standard error of the difference for all t-tests.

For the F-test, the regression equations from the error line and the (Treatments + Error) line must both be solved. If treatments do not affect x or z, an alternative is to begin with an approximate F-test based on the analysis of variance of $(y - b_{y.x}x - b_{y.z}z)$. It will be necessary to compute the exact F-value only in doubtful cases.

9. MORE COMPLEX CLASSIFICATIONS

Three cases that have received attention in the literature are hierarchical classifications, as represented by the split-plot design, incomplete block designs with recovery of interblock information and two-way classifications with unequal numbers in the cells.

In the split-plot design, two independent regression coefficients can be computed, b_1 from the whole-plot error line and b_2 from the sub-plot error line in the analysis of variance. If the two regression coefficients appear to differ, b_1 can be used to adjust whole-plot comparisons and b_2 to adjust sub-plot comparisons. The structure of the analysis of variance is given by Kempthorne [1952]. If the suffix i denotes the whole-plot treatments and j the sub-plot treatments, the adjusted means in the two-way table of treatment means can be computed as follows.

$$y'_{ij.} = y_{ij.} - b_1(x_{i..} - x_{...}) - b_2(x_{ij.} - x_{i..}) \tag{30}$$

Although equation (30) looks as if a bivariate regression is being used to make the adjustments, comparisons between whole-plot treatment means $y'_{i..}$ are actually adjusted by b_1 alone, since

$$y'_{i..} = y_{i..} - b_1(x_{i..} - x_{...})$$

This form of adjustment avoids a discomforting feature mentioned by Bartlett [1937] and Truett and Fairfield Smith [1956], namely that individual adjusted means may not average to the appropriate whole-plot treatment means. Similarly, it is easy to verify that comparisons between sub-plot treatments and components of the interaction between whole-plot and sub-plot treatments are adjusted by b_2 alone. Both regression coefficients enter, however, into certain particular comparisons. For instance, to compare two whole-plot treatments for the same sub-plot treatment, we take, say,

$$\begin{aligned} y'_{2j.} - y'_{1j.} = y_{2j.} - y_{1j.} - b_1(x_{2..} - x_{1..}) \\ - b_2(x_{2j.} - x_{1j.} - x_{2..} + x_{1..}) \end{aligned} \tag{31}$$

As in the split-plot design without covariance, standard errors of comparisons like this require special investigation. For (31), the estimated variance works out as

$$\frac{2}{r}\frac{s_1^2 + (\gamma - 1)s_2^2}{\gamma} + \frac{(x_{2..} - x_{1..})^2 s_1^2}{E_{1xx}} + \frac{(x_{2j.} - x_{1j.} - x_{2..} + x_{1..})^2 s_2^2}{E_{2xx}}$$

where γ is the number of sub-plot treatments, s_1^2 , s_2^2 are the whole- and sub-plot error mean squares, and E_{1xx} , E_{2xx} are the whole- and sub-plot error sums of squares for x.

If the two regression coefficients can be assumed to be the same, a good working procedure, although not fully efficient, is to adjust all means by the sub-plot coefficient b_2 . This procedure, originally suggested by Bartlett [1937], has also been recommended after further examination by Truett and Fairfield Smith [1956].

Similar issues arise with incomplete block designs, since separate regressions can be calculated from the Blocks and Intra-block error lines in the analysis of variance. The computations for recovery of inter-block information, using the intra-block error regression to adjust all means, are illustrated for the 9×9 triple lattice by Cox and Eckhardt [1940] and for the 7×8 simple rectangular lattice by Robinson and Watson [1949]. The approximations involved in this method need further investigation, particularly for the smaller designs.

Methods for handling covariance analyses in the simpler non-orthogonal multiple classifications are given by Wilks [1938], Hazel [1946] and Das [1953], while Federer [1955] presents methods for lattice squares and changeover designs in which residual effects are to be estimated.

REFERENCES

Anderson, R. L. and Bancroft, T. A. [1952] *Statistical theory in research.* McGraw-Hill Book Company, New York. Chapter 14.

Bartlett, M. S. [1937] Some examples of statistical methods of research in agriculture. *J. Roy. Stat. Soc. Suppl. 4:* 137–183.

Cochran, W. G. [1938] Crop estimation and its relation to agricultural meteorology. *J. Roy. Stat. Soc. Suppl. 5:* 12–16.

Cochran, W. G. [1947] Some consequences when the assumptions for the analysis of variance are not satisfied. *Biometrics, 3:* 22–38.

Cochran, W. G. and Cox, G. M. [1957] *Experimental designs.* John Wiley and Sons, New York. 2nd Edition.

Cox, G. M. Eckhardt, R. C. and Cochran, W. G. [1940] The analysis of lattice and triple lattice experiments in corn varietal tests. Iowa Agr. Expt. Sta. Res. Bull. 281.

Das, M. N. [1953] Analysis of covariance in two-way classification with disproportionate cell frequencies. *J. Ind. Soc. Agric. Stat. 5:* 161–178.

Day, B. and Fisher, R. A. [1937] The comparison of variability in populations having unequal means. An example of the analysis of covariance with multiple dependent and independent variates. *Ann. Eug. 7:* 333.

Federer, W. T. and Schlottfeldt, C. S. [1954] The use of covariance to control gradients in experiments. *Biometrics 10:* 282–290.

Federer, W. T. [1955] *Experimental design.* Macmillan Company, New York. Chapter 16.

Finney, D. J. [1946] Standard errors of yields adjusted for regression on an independent measurement. *Biometrics 2:* 53.

Fisher, R. A. [1932] *Statistical methods for research workers.* Oliver and Boyd Ltd., Edinburgh. 4th Edition.

Greenberg, B. G. [1953] The use of analysis of covariance and balancing in analytical surveys. *Am. J. of Pub. Health, 43:* 692–699.

Hazel, L. N. [1946] The covariance analysis of multiple classification tables with unequal subclass numbers. *Biometrics 2:* 21–25.

Kempthorne, O. [1952] *The design and analysis of experiments.* John Wiley and Sons, Inc., New York, p. 387.

Nair, K. R. [1939] The application of covariance technique to field experiments with missing or mixed-up yields. *Sankhya 4:* 581–588.

Robinson, H. F. and Watson, G. S. [1949] Analysis of simple and triple rectangular lattice designs. North Carolina Agr. Exp. Sta. Tech. Bull. 88.

Truett, J. Titus, and Smith, H. F. [1956] Adjustment by covariance and consequent tests of significance in split-plot experiments. *Biometrics 12:* 23–39.

Wilks, S. S. [1938] The analysis of variance and covariance in non-orthogonal data. *Metron. 13:* 141–154.

Yates, F. [1933] The analysis of replicated experiments when the field results are incomplete. *Emp. Journal Exp. Agr., 1:* 129–142.

Yates, F. [1934] A complex pig-feeding experiments. *J. Agri. Sci. 24:* 519.

Yates, F. [1938] Orthogonal functions and tests of significance in the analysis of variance. *Jour. Roy. Stat. Soc. Suppl. 5:* 177–180.

*(P. 263) For a more thorough comparison of covariance and blocking in this situation, see D. R. Cox, Biometrika, 44, 150-158, 1957.

INTERPRETATION OF ADJUSTED TREATMENT MEANS AND REGRESSIONS IN ANALYSIS OF COVARIANCE*

H. Fairfield Smith
Institute of Statistics, North Carolina State College, Raleigh, N.C., U.S.A.

1. INTRODUCTION

Statistical analysis with several variates or variables has become loosely classified into multivariate analysis, regression analysis and analysis of covariance. Most workers have more or less definite ideas about what is implied by each term, but few would agree with any specific definition of their boundaries. Most multivariate work is in one sense or another covariance analysis; but, since its introduction by Fisher [1934] as an adjunct to analysis of variance, the term "analysis of covariance" has had a special connotation, although the restriction may not always be maintained.

Analysis of variance has itself acquired multiple connotations. In line with the comment by Eisenhart [1947] its etymological sense might be taken to imply detection and estimation of components of random variation associated with a composite population—now often referred to as variance component analysis. In a wider sense some have thought it might more accurately have been described as analysis of a sum of squares, particularly when the model specifies only one variance derived from random sources, and the parts of the analysis are compounds of this with constants by which class means are supposed to differ. In this form it is primarily an algorithm for tests of significance for estimates of certain constants, estimands of location. The associated model is conveniently called a regression model because the constants to be estimated can be formulated as regression coefficients. (See, for example, Anderson and Bancroft [1952]). When classifications (treatments) are qualitative the estimands can be regarded as regression

*Part of a review of Analysis of Covariance sponsored by the Office of Ordnance Research, U. S. Army, under Contract DA-36-034-ORD-1517; and contributed to a Symposium on Analysis of Covariance at the Detroit meeting of the Biometric Society (ENAR), American Statistical Association and Institute of Mathematical Statistics, September 7, 1956.

coefficients associated with "dummy variables"; when treatments are quantitative they may be formulated as regression coefficients in the usual sense of average rates of change relative to continuous independent variables. Analysis of covariance for regression estimation pertains especially, although not exclusively, to an extention of the regression model for analysis of variance.

This paper considers two of the more important problems which arise in applications of covariance analysis: the interpretation of adjusted means, and the comparison of treatment ("external") and error ("internal") regressions. Sums of squares and products in analysis of variance and covariance of a dependent variate (y) and a concomitant variable (x) are indicated as follows:

		Sums of Squares and Products		
	D. f.	x^2	xy	y^2
Between treatments	$(t-1)$	T_{xx}	T_{xy}	T_{yy}
Error	ν_E	E_{xx}	E_{xy}	E_{yy} .

The following notation will also be used in the discussion:

$b_T = T_{xy}/T_{xx}$ = treatment regression of y on x
$b_E = E_{xy}/E_{xx}$ = error regression of y on x
$T^* = T_{yy} - b_T T_{xy} = T_{yy} - Y_T^2$ = sum of squares for treatments adjusted by b_T
$E^* = E_{yy} - b_E E_{xy} = E_{yy} - Y_E^2$ = sum of squares for error adjusted by b_E
$Y_T^2 = T_{xx} b_T^2$ = square due to treatment ("external") regression
$Y_E^2 = E_{xx} b_E^2$ = square due to error ("internal") regression
Y_1^2 = square due to regression for pooled treatment and error

$$Y_2^2 = \frac{T_{xx}E_{xx}}{T_{xx} + E_{xx}} (b_T - b_E)^2 = Y_T^2 + Y_E^2 - Y_1^2$$

= square due to "heterogeneity" of treatment and error regressions $Y_3 , \cdots , Y_t = (t-2)$ other linear functions of y such that $\sum_{k=3}^{t} Y_k^2 = T^*$ = sum of squares for deviations of treatment means from treatment regression, and $\sum_{k=2}^{t} Y_k^2$ = reduced treatment sum of squares.

We assume the usual model,

$$y_{ij} = \mu + \tau_i + \beta x_{ij} + \epsilon_{ij} ,$$

where y_{ij} and x_{ij} are the jth observations on the ith treatment of the dependent variate and concomitant variable, respectively, where the

general mean $x_{..}$ is zero, and where the ϵ_{ij} are independent, random errors with zero means. Normality of the ϵ_{ij} is essential only for significance testing and interval estimation. The notation $x_{i.}$ and $y_{i.}$ will be used to indicate treatment means. In general we deal with the case of equal numbers of observations per treatment. The properties of the x_{ij} are not specified above, and this paper is concerned primarily with necessary properties and with what happens if these properties do not hold.

2. THE BASIS FOR USING COVARIANCE TO REDUCE VARIATION

A postulate underlying the application of analysis of covariance to reduce the effect of extraneous variation on estimated treatment responses is that the concomitant variable, x, is unaffected by treatments, either by direct causation or through correlation with another affected character. Given that postulate, means can be adjusted to give more precise estimates of treatment effects than if analysis of covariance were not used and interpretation in the usual way seems unambiguous. The first part of this paper will be mainly concerned with the very different conditions when the postulate may fail. But as a preliminary we will note two matters which are frequently assumed to have a meaning which they do not necessarily have.

Assuming the foregoing postulate to be true, the *formulation of orthogonal linear functions of the variate specified in section* 1 *is a purely formal analysis* for the purpose of proving in an elementary way the properties of the partitions of the analysis of variance. All that is required for this purpose is to define $(t - 1)$ independent linear functions which contain all the information about treatment responses unconfounded with other parameters, in particular β, and having homogeneous error variance. The easiest way to do this is to formulate $(t - 2)$ comparisons among observed means which are independent of β, and a single remaining contrast, Y_T , in which alone confounding with β has to be considered. This last is the regression of treatment means, that is the regression of $y_{i.}$ on $x_{i.}$ which is

$$b_T = Y_T/\sqrt{T_{xx}} = \sum x_{i.}y_{i.}/\sum x_{i.}^2$$

$$= \sum x_{i.}\tau_i/\sum x_{i.}^2 + \beta + \sum x_{i.}\epsilon_{i.}/\sum x_{i.}^2$$

Variation of observed treatment means of the variate, $y_{i.}$, which can be correlated with $x_{i.}$ being concentrated in this single contrast, decontamination of the treatment contrasts is easily seen to be accomplished by subtracting the independent estimator of β which is available from contrasts among experimental units within treatments. Thus the

information about τ_i which is latent in Y_T is salvaged by

$$\sqrt{\frac{(T+E)_{xx}}{T_{xx}E_{xx}}}\, Y_2 = (b_T - b_E) = [\sum x_{i.}\tau_i / \sum x_{i.}^2.]$$

$+$ terms in ϵ_{ij} with expectation zero.

The expression in brackets might be regarded as the regression of τ on x, but it is not usefully so regarded because the postulate affirmed that τ and x are unrelated and therefore that on the average of many experiments that regression must be zero. To test for a relationship which is known not to exist is entirely pointless. The only role of Y_2 is a purely mathematical one of making with $Y_3 \cdots Y_t$ a complete set of $(t-1)$ independent linear contrasts, with equal error variance, and containing all available information on treatment responses. From these any other linear contrast can be compounded, and its error variance indicated in a theoretically elementary manner; but individual contrasts dictated by the $x_{i.}$ which happen by chance to be associated with $y_{i.}$ in a given experiment would rarely be of interest. In practice they are never computed since the sum of their squares can be otherwise obtained, and individual contrasts of special interest can be more directly computed as the required contrast of observed means with a regression adjustment.

When the postulate cannot be made, or if it may be in doubt, Y_2 may acquire some individual interest. But the possibility that $\tau_{i.}$ is associated with $x_{i.}$ by more than chance is *not* tested by comparing Y_2 to internal error. This point will be discussed more fully in sections 7–10.

The concomitant variable is not necessarily a factor causally affecting the variate. The second point, which seems seldom recognized, is that, again assuming the postulate to be true, we need no assumption about the causal relation between x and y. It suffices only that x be correlated with something, often unknown, which causes extraneous variation of y. For example, we might use covariance on amount of a weed present in field plots. But the weed itself may have had no effect on the crop yield; unknown to us associated variation of yield may be due to soil acidity with which growth of the weed is also correlated. Assuming variable acidity to have been a prior soil character independent of treatments, comparison of yields adjusted by the observed regression to a common weed intensity is valid to eliminate from comparisons some of the variability due to acidity and any other incidental circumstances with which weed intensity may be correlated. Control of the weed by cultivation or herbicide will not produce alteration of yield as stated by the regression. Without good ancillary knowledge of causative

influences one should not assume that the character measured as x is responsible for correlated variation of y. A main purpose of this paper is to demonstrate that such assumptions may have resulted in many false interpretations.

Although examples given here are mostly from agriculture the arguments are general. For example in place of plant density an industrial chemist may read impurities in a reagent or still and make obvious transpositions of relevant questions. Was variable contamination accidental? Did it vary between treatments due to different sources of supply? Will it always be associated with respective treatments, or can it and will it be controlled in future? And so on.

A standard example is the effect of plant density on yield of field plots. It may vary between plots due to careless work by the planters, due to action of a drill being faulty or to its seeding rate varying with weight of seed in the hopper, due to variable germination associated with soil effects, or due to attacks by vermin. All of these should be randomly associated with treatments, error and treatment mean squares of x should not differ significantly, and the model is directly applicable.

3. DIFFERENCES OF CONCOMITANT VARIABLE IMPOSED WITH TREATMENTS

If treatments are varieties with different seed sizes, sowing all with a common drill setting may result in different densities per variety. If comparison at common density is appropriate, we could, and perhaps should, try to achieve this by adjusting the drill settings. But whether or not this is attempted, there will always be some variation in the number of seeds actually planted. If one knows that variable plant density within varieties derives from seed numbers planted, one may feel confident that the regression estimates the effect of alterations of seed rate and can be used to estimate what would have been the yields of the varieties at seed rates different from those actually sown, in particular at equal numbers of seed for all. Now suppose that plant number, in addition to being affected by number of seeds planted, is also affected by variable germination associated with soil fertility. In this case the covariance adjustment may be misleading. Fertility has a direct effect on yield quite independent of plant number, whereas plant number by itself may not, in fact often does not, affect yield per area at all. Thus if the number of seeds sown per variety could be controlled exactly, the regression of yield on plant number would in reality be a regression of yield on fertility, and adjustment to a common plant number would be equivalent to adjustment to a common level of fertility. But in practice seed numbers cannot be controlled

exactly, and when they differ enough to affect plant number significantly, a covariance adjustment to a common plant number is in general an adjustment to different levels of fertility -quite contrary to what is intended. This unhappy situation arises because for any given plant number, there are many combinations of seed number and fertility which would produce that particular plant number, and it would be impossible to associate each level of fertility with a single plant number.

If one wants to use an observed regression to allow for an ancillary environmental effect which cannot conveniently be randomized with treatments and may vary systematically from one treatment to another, the only safe procedure is to arrange for it to be deliberately varied. Even when one is confident that a regression is attributable directly to the observed concomitant variable, unaffected by an intermediate correlate, if its variation within treatments be left to chance, it may happen that internal variation is too low either to evaluate the regression accurately or to overlap between treatments; the error variance of consequent extrapolated, or near extrapolated, adjustments may then be excessively high. The ideal arrangement, if an ancillary factor may seriously affect results, is to factorize it with treatments even if its levels may not be precisely controlled. For example Smith [1939] described a spacing experiment with wheat varieties where the field workers failed to plant the same number of seeds in the replicates of each treatment. There was no doubt about the reason for unequal plant numbers between replicates and therefore no doubt that the within and between spacing regressions would be evaluating the same thing. Responses were therefore evaluated as a regression of yield on observed plant numbers irrespective of the nominal spacings (actually several regressions were evaluated, one for each variety). But if spacing replicates had been uniformly sown and variation of plant numbers developed owing to field conditions, the internal and external regressions might, probably would, have been different. Although nominally both would be for yield on plant number they would be used differently, the internal being used to adjust in the usual way for variable soil conditions of unknown origin, and then the regression on imposed variation of plant density evaluated from the adjusted means.

4. CONCOMITANT VARIABLE AFFECTED BY TREATMENTS

The next situation to be considered arises when differences of x between treatments derive, not from imposition with the treatments, but as a consequence of treatments. Varieties, fertilizers, insecticides, etc., may affect germination and hence plant density. With foresight, after preliminary tests, one might have counteracted with different

seed rates. If such action be reasonable, then evaluation of likely results had we done so may be in order; but it may be wise to remember that now the differential germination is a part of the treatment effect, and to consider whether such counteracting action would in fact be taken in practice. Should the treatments be more appropriately compared at equal or at different seed rates? All too often writers blithely assume, without discussion of what "adjusted" means really estimate, that they are "more correct" estimates of treatment effects. More cautious writers seldom go beyond the following remark (Saunders and Rayner [1951] p. 123) appended to an example with unequally germinating varieties of corn: "If variations in stand form an integral portion of the treatment effect, the covariance analysis should not be applied except in so far as it may be desired to assess the influence of treatments apart from any differences in stand which they may cause." But even this reservation may not be enough. If varieties were planted with equal seed numbers, what caused the variation among replicates, and can the consequent regression be assumed to represent what would happen if the poorer germinating varieties were more densely planted? Or did poor germination of some varieties derive from using old seed; and if the counteracting action is to use fresh seed, with greater growing as well as germinating vigor, would the response be anything like that of the internal regression of the given experiment? Suppose in the first example above that fertilizer or spray treatments had affected the quantity of weed, the interpretation put on adjusted yields is too often equivalent to saying that they represent treatment responses to be anticipated if weeds were equalized by cultivation: a possible interpretation but probably false. When the influences causing an observed regression are unknown, estimates obtained by projecting along it, perhaps to an x which is incompatible with a given treatment, have little if any meaning. Frequently they cancel out effects which an experiment was intended to detect.

Snedecor ([1946], Table 12.13) quotes an experiment in which fertilizers affected both germination and yield of beets. The observed regressions for yield *per beet* in arbitrary units on plant number per plot were:

Between fertilizers	(6 d.f.) b =	$.47 \pm .02$
Between replications	(5 d.f.) b =	$-1.16 \pm .43$
remainder	(30 d.f.) b =	$.25 \pm .11$

For a change in plant number only, without associated fertility effects, such a regression should be negative, i.e., increased number of plants per acre results in smaller yield per plant. The remainder regression in

this experiment seems to be a fertility effect and using it to estimate the effect of equalizing stands by altered seed rates would adjust in the wrong direction and be worse than useless. Only rarely will there be so plain a warning that an observed regression does not represent what one might automatically suppose; the hazard for estimating effects of modified treatments must be nearly always present and yet has been almost universally ignored even when, as here, contradiction of a known effect raises a red flag.

In this experiment the observed treatment and error mean squares for plant number (x) and the observed and reduced mean squares for yield *per acre* (y) were

	D.f.	Observed m.s. x	Observed m.s. y	Reduced m.s. (y)
Treatments	6	19337	18.809	.411
Error	30	956	.774	.241

The customary inference is to admit that adjusted means have had some treatment effect eliminated from them, but to say that since the reduced treatment mean square is now insignificant we have obtained the "valuable"* additional information that the treatments have produced nearly all their effect by way of effects on plant number. But this does not follow. Without direct experimental variation of plant numbers we cannot confidently know what effect on yield per area would have been caused by altering them alone. Increasing them would almost certainly depress yield per beet and wide variation around customary seeding rates usually has little effect on yield per area. The most that can be allowed for effect of plant number on yield per acre is to increase it proportionally when the regression for yield per beet is zero. Such response could be expected only when plant densities are so low that plants do not compete with each other, and even if that unlikely condition existed in this experiment it could account for less than half of the observed response to fertilizers.

5. DEPENDENT VARIATE AND CONCOMITANT VARIABLE BOTH CHARACTERS OF THE EXPERIMENTAL MATERIAL

Consider finally the situation when x and y are both measurements of the experimental material. For example, let treatments be varieties of corn, and x the number of ears at constant plant density, a precursor

*Not a quotation from any particular paper; the adjective commonly appears in the context indicated.

of yield. Ear number, an innate variety characteristic, cannot be altered at will. Comparison of yields adjusted to equal ear number is then artificial because we cannot in reality produce equal ear numbers under conditions in which yields would ordinarily be comparable. We might be able to alter ear numbers by changing plant densities or fertilizers, but the comparison adjusted to equal ear numbers would then be a comparison for different, perhaps violently different, environmental conditions.

For the sake of argument suppose that for some reason one is interested in such comparison, say for ear numbers equalized by different densities of stand. If interpretation is to be unambiguous the inference must, as above, be based on a regression which truly reflects the intended mode of variation. Suppose observed plots vary substantially in both fertility and plant number, but that the plant numbers (p) depend only on seed rates which were randomized with plots. Since fertility affects both yield (y) and ear number (e) the regression of y on e is a mixture of fertility and plant number effects, and will not measure the effect of changing e by varying p. Regression on p should not however be affected by fertilities with which it has been randomized. Assume that all regressions are linear, and for simplicity consider the contrast of just two varieties (A, B) with observed means $\bar{y}_A$, etc. From the regression of e on p we may estimate that varieties would have equal e when $p_A - p_B = (\bar{e}_A - \bar{e}_B)/b_{ep}$. And from the regression of y on p estimate that if compared at p_A and p_B the difference of yields would be

$$\hat{y}_A - \hat{y}_B = \bar{y}_A - \bar{y}_B - (\bar{e}_A - \bar{e}_B)b_{yp}/b_{ep} .$$

That is, to answer the question as to whether yields would be equalized by altering p to force equal e we should perhaps compare the ratio $(y_A - y_B)/(e_A - e_B)$, with b_{yp}/b_{ep} , rather than with the observed b_{ye} within treatments. (Compare the use of "instrumental variables." A recent review of the topic is given by Durbin [1954].) Distribution problems with ratios of regression coefficients may be difficult, and field experiments would rarely be sufficiently accurate to yield information on such refined comparisons. We are concerned only to point out that adjustment of one variate for a mutually dependent one by means of a regression derived from variation of uncertain origin leads to estimates without tangible interpretation. Many workers, warned by Bartlett [1936], are aware of the difficulties; but others who present such statistics rarely show recognition of having estimated something completely different from an equalization of extraneous variation.

Two papers written in the first flush of enthusiasm for a new technique and reporting such adjustments (Garner et al [1934]; Brady [1935]) have been much quoted as model examples of how covariance analysis may aid and extend interpretation of experimental data. Comparison of regressions may be informative (cf. sec. 7-10), but many hours endeavoring to see why the adjusted means of these papers are supposed to be helpful invariably end (for me) in a sense of frustration and thinking in circles! When treatments induce simultaneous variation in all characters to estimate their effects on one while "holding another constant" is artificially fictitious. To my mind such estimates are anything but illuminating. The commonly occurring phrase "corrected means," with its implication that more correct comparisons are thereby made, is misleading; "fictitious means" might be less deceptive. In absence of experimental control to demonstrate what in fact would happen if a given character were "held constant," one can seldom demonstrate that actually false conclusions are reached, but treatment means "adjusted" by error regressions seem to yield no information which cannot be gleaned from simpler statistics unconfused by distortions whose interpretation is dubious.

To illustrate let us re-examine Brady's example which is still quoted in standard expositions (e.g. Wishart [1950]). He was studying thickness of sclerenchyma cell wall (y) in oats as a possible index of resistance to lodging, and considered its relation to length (x_1) and diameter (x_2) of an internode. He grew three varieties, which differed markedly in lodging potential, at three spacings in a 9×9 latin square. His analyses of variance, crude and reduced, and variety means, were as in Table 1. His conclusions were:

"In comparing the corrected means, the significant value of z already obtained [for the reduced variety mean square] is undoubtedly due to Sandy being lower than Sonas or Victory II, the difference between the latter two varieties being quite insignificant. It is noteworthy that according to the conclusions following [the analysis of variance of observed y] Sonas was adjudged as having significantly thicker sclerenchyma cell walls than Victory II. It is evident that most of the original mean difference between these varieties is due to the fact that the Sonas plants had shorter internodes and greater internode diameters than the Victory II plants, and that these characters are respectively negatively and positively correlated with thickness of cell wall. It is now possible to make an important deduction, namely, that the conclusions arrived at from the analysis [of variance of observed y] have no general significance. It is known that length of internode is easily varied by changes in soil fertility. Consequently, on

TABLE 1

MEAN SQUARES AND MEANS FOR BRADY'S EXPERIMENT

	D.f.	Mean Squares (y) Observed		Reduced
Variety	2	11.0491		1.1309
Spacing	2	1.0465		.2402
Interaction	4	.0912		.1839
Error	( )	(56) .3482	(54)	.3135

	Variety Means Observed			"Adjusted"
	x_1	x_2	y	y
Sonas	9.13	5.59	5.12	4.79
Victory II	10.56	5.01	4.71	4.72
Sandy	13.45	5.28	3.87	4.19

account of the correlation between length of internode and thickness of cell wall, it follows that the latter would also be subject to variation as a result of soil heterogeneity. With different soil conditions it might be possible to obtain results of an entirely different order from [the observed means]. Hence, thickness of cell wall cannot be taken as providing an absolute index of the loding resistant potentialities of a variety."

The heart of this conclusion lies in one word, "absolute," which would be easily overlooked in hasty reading. The author was merely concluding that the sclerenchyma observed on one variety at any time and place would not be a certain guide for grading it relative to another variety observed in different conditions. This is a reasonable conclusion, and is only what we should expect. Constancy between varieties is possible for some qualitative characters such as petal number and color, but is rare and unlikely for quantitative characters. Even the difference between Mendel's giant and dwarf peas could be masked by growing the one in impoverished, the other in fertile soil. If sclerenchyma might show an equally clear cut genetic effect, only direct experimentation in different conditions will prove how much environmental variation it can stand before varietal differences would be masked. The regression argument provides only circumstantial evidence. Firstly it has not been demonstrated how much change of x_1 and x_2 can be produced by varying environment, and secondly the regression projections are

unreliable as proof of what would happen to y when they are thus varied. The reduced mean square can be misleading as an indicator of what difference in y might still exist when variety-environment combinations were arranged to equalize the x's. It has been literally reduced to be comparable to an error term excluding errors of adjustments. Suppose the x's were equalized by the same influences as operated in the observed latin square. The mean square for variety means adjusted by the observed regression would be 2.949 whereas the reduced mean square was only 1.131. For Sonas-Victory the observed sclerenchyma difference was .41 ± .161 microns; the "adjusted" difference was .084 ± .204. But the adjustment alone (−.32) has a standard error of ±.136 so that with 95 per cent confidence we do not really know but what the correct adjustment might have been only −.05. Corresponding figures for Sonas-Sandy are observed difference 1.25 ± .161; "adjusted" difference .617 ± .275; adjustment −.63 ± .229 with confidence only that it would probably be bigger than −.17. These estimates depend furthermore on the postulate that the appropriate regression has been estimated, a postulate which has been noted above to be often rather shaky. Hence, such projections are not conducive to confident evaluation of what would really happen.

Anyhow regression arguments among mutually variable characters are inappropriate. The speculation is that the internode characters which are easily observable may provide as good evidence of lodging potential as internal characters which might not then be worth using owing to the difficulty of observing them. The speculation is reasonable, but evidence for or against it should derive from a regression analysis of lodging as a variate on all potential predictors studied on equal footing as independent variables. Alternatively, in absence of scaling for lodging resistance, one might consider a canonical variate or principal component analysis for discriminating between varieties. One could, of course, state that two very highly correlated characters (or sets of characters) would give little better discrimination than either one alone. But the multiple correlation of sclerenchyma with both internode characters is only $R^2 = .132$ in the remainder row of the covariance analysis with 56 degrees of freedom, ($P = .02$), or $R^2 = .195$ for total variation within varieties with 78 degrees of freedom including spacing effects. (With only three varieties the between variety correlation of course cannot be evaluated.) These figures therefore do *not* support the conclusion that sclerenchyma may add little or nothing to discrimination.

Conclusions adduced from the covariance analysis of Brady's experiment contain two phrases which commonly occur in reports of

correlation studies. "Most of the original mean difference [of sclerenchyma] between these varieties is due to the fact that" they had different internode lengths and diameters; and "differences between the variety means for thickness of cell wall were exaggerated by reason of this correlation" (Wishart [1950]). Apart from the biological correlations being low with causative connections unknown, which circumstances make the statement even more dubious, such inferences seem analogous to saying that the difference between observed heights of Mt. Everest and Pike's Peak is "due to" air density and is "exaggerated," implying that it is false, because the difference adjusted for correlation of altitude with atmospheric density would be negligible! Obviously such conclusions are incorrectly, or at best metaphorically, phrased; but even after recognizing that, the intended meaning always seems to remain rather puzzling. One may speculate on the consequences for something else—relative difficulties of climbing or assessment of lodging potential -if circumstances promoting the correlation were altered and how. The speculation may be idle curiosity—suppose the summit of Everest were at 10,000 feet; or it may be about something of practical consequence—could Everest be climbed as easily as Pike's Peak after inventing oxygen and pressure suits? The latter type of suggestion, especially when derived from low correlations of unspecified origin should be recognized to be, not a conclusion, but a speculative hypothesis until it can be verified or modified by demonstrating what in fact are the consequences after the environment has been experimentally altered in the specified manner.

6. A PHYSICAL-CHEMICAL EXAMPLE

On rubber plantations hydrometric observations, that is measurements of density, are commonly used to indicate dry rubber content of latex. In an experiment on the effect of some treatments on a physical or chemical character of latex it might be used as a concomitant variable. It may serve well to control effects of variability in both d.r.c. and serum of latices which have been randomized with treatments. But it will not serve as a control for bringing latex from different sources to equal d.r.c. by adding water or ammonia solution; or to estimate results at equal d.r.c. when the treatments impose different degrees of dilution. The imposed variation between treatments would then be on a curve, depending on density of the diluting fluid, which is easily shown (Smith [1940]) to be different from that for the relation between d.r.c. and density of latex as produced by the trees. Furthermore chemical and physical characters are often influenced by serum solids, and whereas a lower d.r.c. with increasing density of field latex implies more serum

solids, to increase density by artificial dilution would reduce these in proportion to rubber content. Hence if, unknown to the experimenter, variability of the observed character happens to depend more on the serum solids than on rubber content, adjustments on density as it varies between latices could be in the wrong direction. Similar remarks apply if the concomitant control were direct observation of rubber content but the variate happens to be affected by serum solids and varying degrees of dilution are introduced with treatments. Within treatments variation derived from different latices would show negative correlation of rubber content and serum solids, whereas between treatments they would be positively correlated and adjustments made on the assumption that rubber content was the sole cause of an observed regression would be false.

7. COMPARING TREATMENT AND ERROR REGRESSIONS

Computationally the partition of sum of squares for treatment regression and mean error regression, Y_T^2 and Y_E^2, into the alternative parts Y_1^2 and Y_2^2 is exactly the same as the analysis of variance for regression heterogeneity. But there is no point to considering Y_2 as a regression contrast $(b_T - b_E)$ unless the between treatment regression has some meaning, which in turn usually implies that we can anticipate or postulate some association between τ_i and x_i, which may be worth estimating. Regression of treatment (or class) responses on x can have meaning only when x has some connection with treatments, either is influenced by them or is a characteristic of them, so that a causal or correlational association of y with x, dependent on treatments may be anticipated. Perhaps the simplest example occurs when treatments are qualitatively determined but have a measurable ingredient variable within them. Two questions may be raised. Can that ingredient be considered responsible for all treatment responses? If it may not account for all responses, is the between treatment response to it the same as that occurring within treatments or is it modified by interaction or by confounding with other ingredients? Perhaps most cases where the external regression has meaning as measure of a quantitative response may formally be reduced by analogy to that form, even although the independent variable is not in ordinary parlance an ingredient.

Boundary cases are worth noting for perspective. When treatments are quantitative—levels of a fertilizer, temperature, concentrations—the responses are efficiently summarized by a regression on the levels, obviously a "between treatment" regression on an "ingredient" which happens here to be the total stimulus applied and to be constant for

replications at each level. We could formally analyze the experiment as an analysis of covariance, but since there is no variation of x within treatments all sums of squares of, and of products with, x would be zero except between treatments, and no internal regression can be evaluated. Hence it would seem artificial so to set out the computations, and without such ado the linear, quadratic, etc., terms of the regression are broken out as special contrasts between treatments.

The ingredient x might be an impurity whose effect we are concerned to "eliminate." There are then two boundary cases. (1) The quantity of impurity may not vary within treatments, and as in the preceding paragraph there is no internal regression. (1a) The treatments proper are qualitative and we can postulate that responses to them will be at random relative to quantities of impurity. The effect of impurities can then be estimated from the external regression and treatment responses by deviations from it. (Notice two incidental consequences. The standard error of the regression on x —which one may wish to examine even if only incidentally—will depend on variance of treatment means about it, not merely on the internal experimental error. The treatment mean square will have only $(t - 2)$ degrees of freedom, corresponding to $Y_3 \cdots Y_t$, and magnitudes of individual responses as measured by deviations from the regression will be proportionately biased.) (1b) Treatments proper are quantitative. Separation will then be by ordinary regression with the treatment means as variates. (2) The impurity is variable within treatments and can be assumed independent of treatment responses. This returns us to the typical analysis of covariance for elimination of extraneous variation by adjusting treatment yields on the internal regression. Although the extraneous variation is in the treatments themselves, instead of as usual in the experimental units, the consequences are obviously the same.

The following sections give some examples where the external regression may be of interest.

8. AN EXAMPLE ILLUSTRATING THE TEST FOR INTERNAL VERSUS EXTERNAL REGRESSIONS

Correlated plant characters in a variety trial or genetic experiment provide an example of internal and external relationships which may be different. Between plots of one variety, strain or clone, (assumed genetically homogeneous), variation is determined by reactions to environmental conditions (fertility, water, etc.); between varieties it is determined by genetic factors. Table 2A presents an analysis of variance and covariance for logarithms of ear number per plant (y) and

of number of grains per ear (x) observed in quintuplicate plots of 9 wheat varieties. (For simplicity some unsatisfactory blocking restrictions are ignored.) Since the environmental variation is of course not entirely averaged out of strain means the estimates of genetic variance components, and by analogy of the covariance component, are obtained as in Table 2B. These suggest a genetic correlation $r = .94$. Should we now assume a functional genetic relation or merely a correlation? If the latter, can a component regression usefully describe the relationship or should some more central line—such as the principal component—be used? Anyway it is at once obvious that the estimate of β_{r_s}

TABLE 2A

ANALYSIS OF COVARIANCE FOR EAR NUMBERS PER PLANT (x) AND GRAIN NUMBERS PER EAR (y) OF 9 WHEAT VARIETIES IN 5 REPLICATIONS

Source of variation	D.f.	Sums of squares and products (for logarithms $\times 10^3$)			
		y^2	yx	x^2	b
Varieties	8	16,329	−16,262	23,724	−.6855
Within varieties	36	5,971	920	14,911	.0617

Source of variation	D.f.	S.s.	M.s.
Variety regression	1	11,147	
Av. env. regression	1	57	
Deviations from var. regression	7	5,182	740.3
Deviations from env. regression	35	5,914	169.0

TABLE 2B

ESTIMATION OF VARIANCE AND COVARIANCE COMPONENTS

Source of variation	y^2	yx	x^2	b
Varieties	2041.1	−2032.8	2965.5	
Within varieties	165.9	25.5	414.2	
Difference = Variance and covariance components $\times$ 5	1875.3	−2058.3	2551.3	−.8068

from $(b_T - b_E)$ as described in section 2 will not be the same as a regression from the variance and covariance components. Even if a regression approach to the problem may be useful models evidently need redefining with some care. Very little work has been done on regression with components (for some suggestions see Tukey [1951]); it is an application of analysis of covariance which extends into multivariate analysis and therefore goes beyond elementary regression formulation and the field of this review.

However to illustrate some other matters, suppose we have a prediction problem, say to predict grain numbers before maturity until when they cannot be observed, and we want to test divergence of regressions for means of 5 plots of different varieties and for single plots within a variety. The varieties are assumed to be a sample of a population of varieties, not necessarily a random sample of the whole population (as the nine here were not, having been selected to represent varieties of different tillering capacity but that each one is a random representative of a sub-population conditional on the observed ear number. (That is of course the elementary condition for an unbiased estimate of any regression.)

Now the precision with which a regression can be determined from a sample depends on the variance of observations about it. Relative to determining a correlation or regression over varieties, the variety means are single observations. Their variance about regression can be equivalent to experimental (environmental or internal) error only if x is a true independent variable free of error and imposes a functional relation on the "true" class means. Hence the precision of the between variety regression must be estimated from $T^*/(t-2)$.

Lest there be any doubt about this, consider that a contrast between just two varieties could be tremendously significant relative to experimental error, but the existence of a correlation cannot be established from only two points it has no degrees of freedom with which to show how much correlation may be relaxed among other class means. Although this should be obvious when pointed out it has not always been recognized. Many fictitious correlations have been reported as if from large numbers of observations when actually they derived from only 3 or 4 dominating class differences.

Since the external and internal regressions are uncorrelated (provided we can assume equal error variance within every class) the estimate of error variance for their difference is

$$\text{Var}\,(b_T - b_E) = \frac{T^*}{(t-2)T_{xx}} + \frac{E^*}{(\nu_E - 1)E_{xx}} = s_d^2 \quad \text{say;}$$

and since its two parts are derived from estimates of variances which cannot be assumed equal the test statistic

$$d = (b_T - b_E)/s_d$$

must be referred to Sukhatmé's table (Fisher and Yates, Table VI) or to the t-table with quasi-degrees of freedom as suggested by Welch [1938], Smith [1936], Satterthwaite [1946] and Cochran [1951]. For the example

$$s_d^2 = \frac{740.3}{23724} + \frac{169.0}{14911} = .04254$$

and

$$d^2 = .7472^2/.04254 = 13.12$$

which is plainly significant without stopping to compute degrees of freedom which must be somewhere between 7 and 42.

Substituting Y_2^2 which may have been extracted in the analysis of variance an equivalent formulation is

$$F' = d^2 = \frac{(T + E)_{xx} Y_2^2}{T_{xx} E_{xx} s_d^2} = \frac{(T + E)_{xx} Y_2^2}{E_{xx} v_T + T_{xx} v_E}$$

where v_T and v_E are the mean squares for deviations about external and internal regressions with $(t - 2)$ and $(\nu_E - 1)$ degrees of freedom respectively.

9. AN EXAMPLE OF A TREATMENT-INDUCED CONCOMITANT

This example, from data supplied by Dr. D. D. Mason, illustrates occurrence of the concomitant variable as a treatment ingredient in fact instead of merely formally. Alfalfa was prepared in three ways—as field-cured hay, barn-cured hay and wilted silage; and was fed to six blocks of three cows each in a switch-back experiment, i.e., the "plots" within blocks are three periods within a three-cow lactation term. Some restrictions on randomization for assessment of carry-over effects are ignored here for simplicity. The observations were X = mgm. of carotene intake in each plot (estimated from analytical samples from each ration), and Y = vitamin A potency per pound of butter fat. The objectives were stated to be: (1) To determine the overall effect of different methods of curing alfalfa on the vitamin A potency of milk, and (2) To determine the effect of the methods of curing independent of the carotene content. The first is answered by the observed means and their standard errors. The purpose here is to give critical

consideration to just how much about the second question can be inferred from covariance analysis.

Logarithms of the data and the analysis of variance and covariance are given in Table 3. The negative correlation between blocks ($r = -.203$) is noteworthy since it masks correlation of the variates within treatments if block effects were not extracted. Although not significant with so few observations, the correlation is probably real due to the units observed; it suggests as is reasonable that large butter producers

TABLE 3

CAROTENE INTAKE (X) AND VITAMIN A POTENCY PER POUND OF BUTTER FAT (Y) FOR THREE TYPES OF ALFALFA FORAGE: $x = 10^3 (\log_{10} X - 2)$, $y = 10^3 (\log_{10} Y - 3.8)$

Block	Field hay		Barn hay		Silage	
	x	y	x	y	x	y
1	149	136	297	138	781	231
2	164	154	393	174	766	280
3	64	111	299	131	675	219
4	90	96	279	207	659	261
5	218	76	389	66	919	263
6	193	85	360	85	979	272
Total	878	658	2017	801	4779	1526
Mean	146.3	109.7	336.2	133.5	796.5	254.3

ANALYSIS OF VARIANCE AND COVARIANCE

	D.f.	Sums of squares & prod. x^2	xy	y^2	Regression b	Square	D.f.	Deviations
Blocks	5	83580	−13437	9886	—	—	—	—
Treatments	2	1341320	308411	72181	.22993	70913	1	1268
Remainder	10	29930	11922	11578	.39833	4749	9	6829
Treatm. + Rem.	12	1371250	320333	83759	.23361	74832	11	8927

	D.f.	M.s.
Heterogeneity of regression	1	830
Deviations from treatm. regr.	1	1268
Reduced treatm. sum of squares	2	2098
Error mean square	9	758 8

eat more but not in proportion to production, total intake is increased but is less per pound of butter.

The analysis of variance looks pleasantly clear cut. Regression squares are high, indicating that carotene content of forage is associated with vitamin A in butter, and both regression heterogeneity and deviations from regression squares are similar to the estimate of error variance. Clearly there is no evidence against the hypothesis that all observations are collinear and that forage effects may derive entirely from their carotene contents. This of course is very far from proving that carotene is the only substance in the forages affecting vitamin A response. For one thing, since all forages are derived from a similar substrate, alfalfa (presumably all from one field), there may be a correlated substance which has not been measured. Apart from that, means adjusted on the internal regression are virtually useless; they involve violent extrapolation and insignificance of their differences means nothing. For example the observed difference of y between silage and field hay is 145 ± 19.6; the adjusted difference is -114 ± 105 with almost all its standard error derived from uncertainty about the regression. The variance is compounded as follows:

$$s^2\left\{\frac{2}{r} + \frac{(\bar{x}_3 - \bar{x}_1)^2}{\text{Error s.s. (x)}}\right\} = 758.8(0.33 + 14.12) = 10970.$$

(Arithmetic data show similar significance tests but different relative magnitudes. The respective differences are observed 3248 ± 410 mgm, adjusted 958 ± 890.) If one were seriously concerned to determine whether the method of curing affects vitamin production other than through its effect on carotene in the forage the experiment needs to be supplemented by hays reinforced by extra carotene so that more direct comparisons can be made. (If the effect of carotene is accurately known from other sources, a better projection might be made; but with the need to make allowances for quantities eaten and produced such theoretical projections for biological material are often doubtful.)

However it is not unlikely that when doing an experiment of this sort one has a fairly confident anticipation that all response will be related to carotene, that results can be summarized in terms thereof, and that objective (2) as stated above is less an objective than a check for any contrary evidence before going on to such summary. The experiment might even have been done primarily to evaluate the effect of carotene as "naturally" varied, perhaps for comparison with other experiments using a "synthetic" form. Clearly the effect may be better evaluated from between treatments owing to the extended range, and this may yet be so even if treatments have a superimposed

qualitative effect if we can assume such to occur at random relative to the mean carotene contents. We will of course check for any disagreement between internal and external regressions, but since in this case the internal is poorly determined with 95 per cent confidence interval .038 to .758 the check does not mean much. The main difficulty if we allow the possibility of superimposed treatment effects is that with only three treatments we have no satisfactory estimate of error for the external regression. Technically its significance stands to be tested by $t = (70913/1268)^{1/2} = 7.62$ with only one degree of freedom and so needing 12.71 to reach the 5 per cent significance level, but since we expect an effect approximately as indicated our subjective confidence is high. There is little however that can be done unless we are prepared to accept the hypothesis that carotene is the only operative agent, and that treatments add no extra responses. The situation then becomes similar to the wheat seed spacing experiment quoted in sec. 3 and the regression can be estimated from total variation within blocks. This yields the estimate $b = .2336 \pm .0243$. Even allowing that the standard error may be under-estimated this should be much better than the internal estimate with error $\pm.159$. (The alternative is to take the treatment regression, .230 with error $\pm.028$ based on the reduced treatment mean square with 2 degrees of freedom. Even with so few degrees of freedom this would still be better than the internal regression.)

Finally one may note that since X can only be estimated from samples extracted from the forage actually eaten it must contain error which could bias the regression as estimate of true response. In a case like this the functional line could be estimated using treatments as "instrumental variate" leading in effect to the Wald-Bartlett-Nair-Shrivastava method of averages; but this refinement is trivial relative to ignorance of the true form of curve which should be fitted.

10. NOTES ON BIBLIOGRAPHY

One or two text books earlier, and some half-dozen since 1950*, have presented expositions of analysis of variance for testing regression heterogeneity including the external versus internal regression, but interpretation of the latter is usually either side stepped or at best

*Among texts on general experimental statistics the test for heterogeneity of regressions in independent samples is given by Villars [1951], Goulden [1952], Tippett [1952]; the break-out of Y_2 for $b_T - b_E$ is given by Wishart and Sanders [1936], and Saunders and Rayner [1950]; both are given by Snedecor [1946], Wishart [1950], Quenouille [1952], Bennett and Franklin [1954]. Of the last two groups only Quenouille seems to recognize that Y_2 is not usually to be tested against internal error. I am not acquainted with sociology texts.

controversial. This is perhaps due to absence of good examples where it is relevant, but may also be attributable to unfortunate examples used by the originators and leading expositors of this feature of analysis of covariance. Better examples are still needed. (The ones used above are not yet good enough!)

(1) E. S. Pearson [1934] first formulated the complete breakdown of the sum of squares. Its application (with computations by Miss F. E. James on data presented by Wilsdon) concerned the regressions of strengths of concretes on strengths of mortars made from the same cements. Unresolved difficulties in both design and analysis may now be recognized.

Welch [1935] simultaneously used likelihood ratio methods to derive the analysis of tests for collinearity, and illustrated on the same data.

(2) The contrast of external versus internal regression was next presented by Wishart and Sanders [1936] with a clarity of algebraic exposition which probably cannot be improved. Their two examples (a theoretical one represented by their figure 2, and one on live data) consider variates which are both affected by treatments; comparison of the two regressions may therefore be relevant. But the discussion purports to represent ordinary covariance usage to eliminate from estimated treatment means some of the random variation associated with an extraneous variable. It shows no recognition of the circumstance that when covariance is used in this way the regression of treatment means on x is irrelevant, and conversely that when both variates are affected by treatments and it is their relationship which matters the adjusted means are artifacts.

(3) Kendall ([1955] sec. 24.30) gives what is perhaps the best available algebraic exposition of the complete regression heterogeneity analysis of hierarchical classification, closely following Pearson [1934] but with a simpler and more straightforward example. (Any edition may be used for the mathematical exposition but the 1955 impression should be consulted for interpretation. A mischievous Puck seems to have played with the earlier text to mate systematic and random effects and accidentally invert ratios—mischances obvious to experts but worrying to students.) However the following ambiguities, more or less latent in Pearson's paper which is being followed, still remain.

Kendall follows the practice of those who take a likelihood ratio approach to multivariate tests for heterogeneity (as illustrated by Welch [1935] and Pearson and Wilks [1933]) whereby one first does a pooled test and then subdivides if required. This raises the thorny problem "to pool or not to pool?" and gets everyone a little confused.

For example the theoretical exposition in effect says, if the pooled test is insignificant consider next the details; whereas the sample says, the pooled test is significant so go on. This seems typical of the progression in multivariate tests! It seems to me one were better to decide in advance the points at which a null hypothesis may be expected to fail, go directly there, and stand by the result. For example, if the analysis of covariance is being done for control of extraneous variability the external versus internal regression heterogeneity is irrelevant, and heterogeneity of internal regressions is best left in as part of random error. If the experiment has been duly randomized in design, so have the consequences of any regression heterogeneity and attempts to take it out get one into supplementary difficulties: to consider regression differences is to go on to something the experiment was not designed to explore and for which it is unlikely to be efficient.

Secondly, Kendall implies that the comparative efficiency of x as a predictor for another character in different groups of individuals can be judged from the magnitudes of the regression coefficients in each group. But efficiency of prediction is independent of slope of regression. It is determined by the precision of prediction—that is in general by variance of the predictand about the regression. (For special cases one may consider also precision with which regressions are estimated from a given size of sample, which depends on dispersion of x, as well as on variance about regression, but is still independent of slope.)

Finally Kendall states that $S_4 (Y_2^2$ in this paper) should be compared to internal error, whereas we have seen that either it is irrelevant, or that when it has meaning the error of b_T , and hence also of Y_2 , depends on deviations about the external regression.

Models: No work seems to have been done on constructing mathematical models for situations where a between treatment regression may be of interest and different from an internal regression. The task seems to be less simple than might be anticipated. A suitable form may need to be derived for each case individually with attention to the physical or biological circumstances.

Unpleasant complications can arise in ill-planned experiments. For example in the urea clearance work discussed by Smith [1951] men were fed supplementary urea quite erratically; and long term changes to reduced protein diets were confounded with renal defect. Consequently variation between men introduced almost inextricable confusion of responses corresponding to natural differences between men, to a day or two of supplemantary urea in diet, to long term diet effects, and to ordinary day-to-day fluctuations.

If an independent variable, x, were controllable its variation could

be imposed as a cross treatment. Comparisons of external and internal regressions seem to become relevant only when x varies within treatments essentially as a random variable and must be so regarded. This makes complications for evaluating expectations of functions of observations and perhaps only criteria of consistency can be handled in elementary manner. Some tentative endeavors to construct models for the foregoing examples seem as yet too crude to be worth presenting at this time.

ACKNOWLEDGEMENT

I wish to thank Messrs. J. A. Rigney, O. Kempthorne and A. W. Kimball for helpful comments.

BIBLIOGRAPHY

A complete bibliography would have to cover regression analysis and innumerable examples of applications. The following is an arbitrary selection which may adequately indicate the development of analysis of covariance in statistical literature.

1. Books

Anderson, R. L. and Bancroft, T. A. [1952] *Statistical theory in research.* McGraw-Hill Book Company, Inc., New York.

Bennett, C. A. and Franklin, N. L. [1954] *Statistical analysis in chemistry and the chemical industry.* John Wiley and Sons, Inc., New York.

Cochran, W. G. and Cox, G. M. [1950] *Experimental designs* (Sec. 3.8) John Wiley and Sons, Inc., New York.

Federer, W. T. [1955] *Experimental designs.* Macmillan Company, New York. (Appendix pp. 46–47 cites references to special experimental designs.)

Fisher, R. A. [1934] *Statistical methods.* Oliver and Boyd Ltd., Edinburgh. 5th Edition.

Goulden, C. H. [1952] *Methods of Statistical analysis.* John Wiley and Sons, New York.

Kempthorne, O. [1952] *Design and analysis of experiments.* John Wiley and Sons, New York.

Kendall, M. G. [1955] *Advanced theory of statistics.* II. 3rd Edition, 2nd impression (Sec. 24.30). Charles Griffin and Co., Ltd., London.

Quenouille, M. [1952] *Associated measurements.* Butterworths Scientific Publications, London.

Saunders, A. R. and Rayner, A. A. [1951] *Statistical methods with special reference to field experiments.* Union of South Africa, Department of Agricultural Science, Bull. 200, 3rd Edition.

Snedecor, G. W. [1946] *Statistical methods.* Iowa State College Press, Ames, Iowa. 4th Edition.

Tippett, L. H. C. [1952] *The methods of statistics.* John Wiley and Sons, New York.

Villars, D. S. [1951] *Statistical design and analysis of experiments for development research.* W. C. Brown and Co., Dubuque, Iowa.

Wishart, J. and Sanders, H. G. [1936] *Principles and practice of field experimentation.* Empire Cotton Growing Corporation, London.

2. Papers

Bartlett, M. S. [1933] On the theory of statistical regression. *Proc. Roy. Soc. Edinb. 53:* 260–283.

Bartlett, M. S. [1934a] The problem in statistics of testing several variances. *Proc. Camb. Phil. Soc. 30:* 164–169.

Bartlett, M. S. [1934b] The vector representation of a sample. *Proc. Camb. Phil. Soc. 30:* 327–340.

Bartlett, M. S. [1935] An examination of the value of covariance in dairy cow nutrition experiments. *J. Agr. Sci. 25:* 238–244.

Bartlett, M. S. [1936] A note on the analysis of covariance. *J. Agr. Sci. 26:* 488–491.

Bartlett, M. S. [1937] Some examples of statistical methods of research in agriculture. *J. Roy. Stat. Soc. Suppl. 4:* 137–183.

Brady, J. [1935] A biological application of the analysis of covariance. *J. Roy. Stat. Soc. Suppl. 2:* 99–106.

Carter, A. H. [1949] The estimation and comparison of residual regressions where there are two or more related sets of observations. *Biom. 36:* 26–46.

Cochran, W. G. [1934] The distribution of quadratic forms in a normal system, with applications to the analysis of covariance. *Proc. Camb. Phil. Soc. 30:* 178–191.

Cochran, W. G. [1948] Analysis of covariance. *Instit. of Stat. Mimeo. Series No. 6.*

Cochran, W. G. and Bliss, C. I. [1948] Discriminant functions with covariance. *Ann. Math. Stat. 19:* 151–176.

Cornish, E. A. [1940] Analysis of covariance in quasi-factorial designs. *Ann. Eug. 10:* 269–279.

Das, M. N. [1953] Analysis of covariance in two-way classification with disproportionate cell frequencies. *J. Ind. Soc. Agric. Stat. 5:* 161–178.

Day, B. and Fisher, R. A. [1937] The comparison of variability in populations having unequal means. An example of the analysis of covariance with multiple dependent and independent variates. *Ann. Eug. 7:* 333–348.

DeLury, D. B. [1946] The analysis of covariance. *Biometrics 4:* 153–170.

Federer, W. T. and Schlottfeldt, C. S. [1954] The use of covariance to control gradients in experiments. *Biometrics 10:* 282–290.

Finney, D. J. [1946] Standard errors of yields adjusted for regression on an independent measurement. *Biometrics 2:* 53–55.

Fisher, R. A. [1947] The analysis of covariance method for the relation between a part and the whole. *Biometrics 3:* 65–67.

Fraser, D. A. S. [1953] The Behrens-Fisher problem for regression coefficients. *Ann. Math. Stat. 24:* 390–402.

Garner, F. H., Grantham, J. and Sanders, H. G. [1934] The value of covariance in analysing field experimental data. *J. Agric. Sci. 24:* 250–259.

Gourlay, N. [1953] Covariance analysis and its applications in psychological research. *Brit. J. Psych. Stat. Soc. 6:* 25–34.

Hazel, L. N. [1946] The covariance analysis of multiple classification tables with unequal subclass numbers. *Biometrics 2:* 21–25.

Hendricks, Walter A. [1935] The use of "differential regression" in analysis of variance. *J. Agric. Sci. 25:* 258–263.

Jackson, R. W. B. [1940] Application of analysis of variance and covariance method to educational problems. *Bulletin XI. Dept. Educ. Res. Univ. of Toronto.*

Kimball, B. F. [1953] A multiple group least squares problem and the significance of the associated orthogonal polynomials. *J. Amer. Stat. Assoc. 48:* 320–335.

Kolodziejczyk, S. [1935] On an important class of statistical hypotheses. *Biom. 27:* 161–190.

Mahoney, C. H. and Baten, W. D. [1939] The use of the analysis of covariance and its limitation in the adjustment of yields based upon stand irregularities. *J. Agric. Res. 58:* 317–328.

Nair, K. R. [1939] The application of covariance technique to field experiments with missing or mixed-up yields. *Sankhyā 4:* 581–588.

O'Neil, J. B. and Gutteridge, H. S. [1941] A note on calculation of standard errors for treatment means after adjusting for regression. *Sci. Agric. 21:* 558–559.

Outhwaite, Anne D. and Rutherford, A. [1955] Covariance analysis as an alternative to stratification in the control of gradients. *Biometrics 11:* 431–440.

Parker, E. R. [1941] Adjustment of yields in an experiment with orange trees. *Proc. Amer. Soc. Hort. Sci. 41:* 23–33.

Pasternack, B. S. [1956] General theory of the analysis of covariance. (unpublished).

Pearson, E. S. [1934] Appendix to paper by B. H. Wilsdon. *J. Roy. Stat. Soc. Suppl. 1:* 178–192.

Quenouille, M. H. [1948] The analysis of covariance and non-orthogonal comparisons. *Biometrics 4:* 240–246.

Quenouille, M. H. [1950] Multivariate experimentation. *Biometrics 6:* 303–319.

Smith, H. Fairfield [1939] Effect of spacing and time of sowing on yield and yield components of wheat varieties. *Counc. Sci. and Ind. Res.* (Australia) *Pamphlet 91.* (also *J. Roy. Stat. Soc. Suppl. 4:* 177–178 [1937]).

Smith, H. F. [1950] Error variance of treatment contrasts in an experiment with missing observations. *J. Ind. Soc. Agric. Stat. 2:* 111–124.

Smith, H. F. [1951] Relationships among characters observed in urea clearance tests. *Biometrics 7:* 185–199.

Truett, J. Titus, and Smith, H. F. [1956] Adjustment by covariance and consequent tests of significance in split-plot experiments. *Biometrics 12:* 23–39.

Tukey, J. W. [1949] Dyadic anova; an analysis of variance for vectors. *Human Biol. 21:* 65–110.

Tukey, J. W. [1951] Components in regression. *Biometrics 7:* 33–69.

Watson, S. S. and Ferguson, W. S. [1936] (Fodder in dairy cow diets, etc.) *J. Agric. Sci. 26:* 337–367.

Welch, B. L. [1935] Some problems in the analysis of regression among k samples of two variables. *Biometrika 27:* 145–160.

Wilks, S. S. [1938] The analysis of variance and covariance in non-orthogonal data. *Metron. 13:* 141–154.

Wishart, J. [1936] Tests of significance in analysis of covariance. *J. Roy. Stat. Soc. Suppl. 3:* 79–82.

Wishart, J. [1939] Statistical treatment of animal experiments. *J. Roy. Stat. Soc. Suppl. 6:* 1–22.

Wishart, J. [1950] Field trials, Part II. The analysis of covariance. *Commw. Bur. Plant Breeding and Genetics.* Tech. Comm. No. 15.

Yates, F. [1934] A complex pig-feeding experiment. *J. Agri. Sci. 24:* 511–531.

Yates, F. [1938] Tests of significance of the differences between regression coefficients derived from two sets of correlated variates. *Proc. Roy. Soc. Edinb. 59:* 184–194.

3. Other Literature Cited

Bartlett, M. S. [1949] Fitting a straight line when both variables are subject to error. *Biometrics 5:* 207–212.

Cochran, W. G. [1951] Testing a linear relation among variances. *Biometrics 7:* 17–32.

Durbin, J. [1954] Errors in variables. *Rev. Inst. Intern. Stat. 22:* 23–32.

Eisenhart, C. [1947] The assumptions underlying the analysis of variance. *Biometrics 3:* 1–21.

Pearson, E. S. and Wilks, S. S. [1933] Methods of statistical analysis appropriate for k samples of two variables. *Biom. 25:* 353–378.

Satterthwaite, F. E. [1946] An approximate distribution of estimates of variance components. *Biometrics Bull. 2:* 110–114.

Smith, H. F. [1936] The problem of comparing the results of two experiments with unequal errors. *J.C.S. and I.R. (Australia) 9:* 211–212.

Smith, H. F. [1940] Specific gravity of latex and of rubber. *J. Rubber Res. Inst. Malaya 9:* 218–247.

Welch, B. L. [1938] The significance of the difference between two means when the population variances are unequal. *Biom. 29:* 350–361.

THE ANALYSIS OF COVARIANCE FOR INCOMPLETE BLOCK DESIGNS*

MARVIN ZELEN

National Bureau of Standards, Washington, D.C., U.S.A.

1. INTRODUCTION

In any experiment the careful experimenter is always trying to "control" all factors which influence the final results of the experiment. In many experiments, however, it is not practical to control all factors even though these factors can be measured quantitatively. For example, in calibrating meter-bars, the temperature of the chamber where the measurements are taken cannot be held constant for repeat measurements, thus necessitating a temperature correction to the measurement. In other experimental situations the innate capacity of the experimental units to respond to the treatments applied to them, is so highly variable that small or even moderate differential treatment responses may be completely obscured. This is the case (say) in applying treatments to used storage batteries where the condition of every cell is highly variable in its ability to respond to a charge. In still other experiments a drift in the measuring equipment is present which may not have been taken into account in the experimental design.

All of the above experimental problems can be handled most effectively by the analysis of covariance technique which was first introduced by R. A. Fisher [13, 5th edition] in 1934. Thus, in the meter-bar measurements, the corrections applied can be made to depend only on the results of the particular measurements themselves, rather than on temperature corrections obtained independently.** The sensitivity of the battery experiment can be markedly improved by applying a

*Paper presented at the joint meeting of the Biometric Society and the American Statistical Association, Detroit, Michigan, September 1956.

**In practice the range of temperatures involved is so minute that the correction so obtained would itself be very imprecise, and recourse is taken to corrections derived from special determinations involving a wide range of temperatures. In those cases where the temperature range is relatively large or when there is a lag between the temperature of the meter bars and that of the chamber, the analysis of covariance can be used to make the adjustment for temperature.

preliminary small charge to each cell before any treatments are applied, then discharging each cell and using this preliminary discharge data to adjust the final test results by taking into account the initial ability of each cell to respond to a charge. The effects of drift in an instrument can, in many cases, be approximated by a polynomial expression of time, thus allowing measurements to be "corrected" for drift.

Covariance analysis can also be used for analyzing experiments when through unforeseen circumstances several observations are missing. If only one observation is missing, or if a very simple experimental design is used, the analysis may not be unduly complicated; however for the more complicated designs the absence of measurements presents a major problem in the analysis For these cases, covariance analysis can be readily applied, resulting in a relatively simple analysis. The procedure first suggested by Bartlett [3] is to let the value of the missing measurements be zero, and introduce concomitant variates corresponding to each missing observation such that the value of the concomitant variate is +1 for the missing unit and zero elsewhere. The number of concomitant variates will be equal to the number of missing experimental units.

Still another use for covariance analysis as pointed out by Bartlett [2] is to adjust treatment effects for systematic differences between experimental units in a non-random experiment. In this situation if the effects of treatment differences disappear after adjusting for initial differences, then one can conclude that the treatments do not differ. However, if after adjusting for initial differences among experimental units the treatment differences still remain, then the conclusion that the treatments actually differ is not necessarily a valid conclusion. Although an exact functional relationship between the dependent and independent variates may not be known, an approximate relationship may be sufficient to adjust for the effects of non-randomness. Thus, if differences among treatments do *not* disappear after making such an adjustment, then we may conclude (i) treatments actually differ or (ii) treatments do not differ, but the functional relationship is not known to a sufficient approximation to adjust adequately for the non-random nature of the experiment. The conclusion that treatments actually differ will only be valid if the explicit form of the adjustment makes use of a known functional relationship between the dependent and independent variates.

In the simplest covariance situation the experimenter applies treatments to experimental units and takes measurements (denoted by y) which he wishes to analyze. However, associated with each experimental unit is an uncontrolled factor, sometimes referred to as a con-

comitant variate, which can be measured or observed at a value x. The value x is completely independent of the treatments subsequently applied, but does reflect the variation among the experimental units with respect to uncontrolled factors. If the response y of the experimental unit to a treatment is affected by the particular value x associated with an uncontrolled factor, then it would be to the experimenter's advantage to "adjust" the measurement for the uncontrolled factor. This "adjustment" *may* take the form of a linear correction to the treatment estimates. That is, if the measurement y varies linearly with x, then $y - Bx$ is the adjusted value for the measurement, where B is a constant which can be estimated *from the present data* and need not depend on other measurements.

Many of the published applications of covariance analysis have dealt with experimental situations where the measurements conformed to a completely randomized design or to randomized block designs, cf. Anderson and Bancroft [1], Snedecor [18], Wishart [21]. The use of the analysis of covariance applied to split-plot experiment designs has been discussed by Bartlett [3], Cochran [7], Kempthorne [15, section 19.7], and Truett and Smith [19]. A monograph by Cochran, Cox, and Eckhardt [8] gives both the inter- and intra-block covariance analysis for triple lattices. Cornish [9] has discussed the intra-block analysis applied to balanced incomplete block designs. Federer [11] has summarized the use of covariance methods applied to completely randomized and randomized block designs, Latin square designs, split plot and split block designs, and various lattice designs. The book by Rao [17] derives the general equations for covariance which are applicable to any experiment design.

This paper, in Section 2, reviews briefly the procedure for the intra-block analysis of covariance when applied to any type of experimental design for p concomitant variates. Section 3 deals with the recovery of inter-block information; section 4 gives formulae for application to balanced and partially balanced (2 associate classes) incomplete block designs; section 5 treats a more general case of covariance analysis where the adjustment for the treatment response depends on the differential block response; and finally section 6 consists of a numerical example illustrating the covariance analysis computations for a balanced incomplete block design.

2. INTRA-BLOCK ANALYSIS

In many of the expositions and applications of the analysis of covariance it "appears" that different formulae are needed for different experiment designs. This is only illusory. Actually, the formulae for

applying the analysis of covariance to various experiment designs is the same for all designs. The results in this section hold for any experiment design or two-way non-orthogonal classification and are a straightforward consequence of the general linear hypothesis. These results were first obtained by Rao [17] and are included for completeness.

Consider an experimental plan having v treatments, arranged in b blocks, such that each block contains k experimental units and each treatment occurs in r blocks. Let the measurement for the ith treatment in the jth block be y_{ij}, and let there be p concomitant variates denoted by $x_{ij}^{(w)}$ $(w = 1, 2, \cdots, p)$ associated with each measurement.

Then the mathematical model underlying the intra-block analysis is

$$y_{ij} = \mu + t_i + b_j + \sum_{w=1}^{p} B_w x_{ij}^{(w)} + \epsilon_{ij} \tag{2.1}$$

where μ is a constant common to all measurements, t_i is the systematic effect for the ith treatment, b_j is the systematic effect for the jth block, B_w is the regression of y_{ij} on $x_{ij}^{(w)}$ for $(w = 1, 2, \cdots, p)$ and ϵ_{ij} is a random variable having zero expectation and constant but unknown variance σ^2. In making all tests of significance it will be further assumed that the ϵ_{ij} are independent identically distributed random variables following a normal distribution.

2.1 *Notation*: It will be convenient to adopt the following notation where it will be understood that a summation with respect to i ranges from $i = 1, 2, \cdots, v$; a summation with respect to j ranges from $j = 1, 2, \cdots, b$; and a summation with respect to u or w ranges from $u, w = 1, 2, \cdots, p$.

Let

$$n_{ij} = \begin{cases} 1 \text{ if treatment } i \text{ appears in the } j\text{th block} \\ 0 \text{ otherwise} \end{cases}$$

$$\begin{cases} Y_{i.} = \sum_j n_{ij} y_{ij}\,, \quad \bar{y}_{i.} = \dfrac{Y_{i.}}{r} & \text{(treatment total and average)} \\ Y_{.j} = \sum_i n_{ij} y_{ij}\,, \quad \bar{y}_{.j} = \dfrac{Y_{.j}}{k} & \text{(block total and average)} \\ Y_{..} = \sum_{i,j} n_{ij} y_{ij}\,, \quad \bar{y}_{..} = \dfrac{Y_{..}}{bk} & \text{(grand total and average)} \end{cases} \tag{2.2}$$

and similar expressions for $X_{i.}^{(w)}$, $X_{.j}^{(w)}$, $X_{..}^{(w)}$, $\bar{x}_{i.}^{(w)}$, $\bar{x}_{.j}^{(w)}$, $\bar{x}_{..}^{(w)}$.

Also let

$$\begin{cases} Q_{i0} = Y_{i.} - \sum_j n_{ij} \bar{y}_{.j} \\ Q_{iw} = X_{i.}^{(w)} - \sum_j n_{ij} \bar{x}_{.j}^{(w)} \end{cases} \text{ adjusted treatment totals, (treatment totals adjusted by block means)} \tag{2.3}$$

$\hat{t}_i$ = estimate for the ith treatment obtained from intra-block analysis;
$\hat{t}_{i0}$ = estimate for the ith treatment obtained from intra-block analysis if $B_w = 0$ for $w = 1, 2, \cdots, p$;
$\hat{t}_{iw}$ = quantity analogous to $\hat{t}_{i0}$, obtained by using $x_{ij}^{(w)}$ in place of y_{ij}.

$$(2.4)\quad \left.\begin{aligned} T_{00} &= \sum_i \hat{t}_{i0}Q_{i0} \\ T_{0w} &= \sum_i \hat{t}_{i0}Q_{iw} = \sum_i \hat{t}_{iw}Q_{i0} \\ T_{uw} &= \sum_i \hat{t}_{iu}Q_{iw} = \sum_i \hat{t}_{iu}Q_{iw} \end{aligned}\right\} \text{adjusted treatment sum of squares and products.}$$

$$(2.5)\quad \left.\begin{aligned} b_{00} &= \sum_j \frac{Y_{\cdot j}^2}{k} - \frac{Y_{\cdot\cdot}^2}{bk} \\ b_{0u} &= \sum_j \frac{Y_{\cdot j}X_{\cdot j}^{(u)}}{k} - \frac{X_{\cdot\cdot}^{(u)}Y_{\cdot\cdot}}{bk} \\ b_{uw} &= \sum_j \frac{X_{\cdot j}^{(u)}X_{\cdot j}^{(w)}}{k} - \frac{X_{\cdot\cdot}^{(u)}X_{\cdot\cdot}^{(w)}}{bk} \end{aligned}\right\} \text{unadjusted block sum of squares and products}$$

$$(2.6)\quad \left.\begin{aligned} S_{00} &= \sum_{i,j} n_{ij}y_{ij}^2 - \frac{Y_{\cdot\cdot}^2}{bk} \\ S_{0w} &= \sum_{i,j} n_{ij}x_{ij}^{(w)}y_{ij} - \frac{X_{\cdot\cdot}^{(w)}Y_{\cdot\cdot}}{bk} \\ S_{uw} &= \sum_{i,j} n_{ij}x_{ij}^{(u)}x_{ij}^{(w)} - \frac{X_{\cdot\cdot}^{(u)}X_{\cdot\cdot}^{(w)}}{bk} \end{aligned}\right\} \text{total sum of squares and products}$$

$$(2.7)\quad \left.\begin{aligned} E_{00} &= S_{00} - b_{00} - T_{00} \\ E_{0w} &= S_{0w} - b_{0w} - T_{0w} \\ E_{uw} &= S_{uw} - b_{uw} - T_{uw} \end{aligned}\right\} \text{error sum of squares and products}$$

2.2 *Analysis*: The estimates of the unknown parameters μ, t_i, b_j, B_w, can be found by minimizing the sum of squares of the errors $\sum_{i,j} n_{ij}\epsilon_{ij}^2$ with respect to each parameter. On differentiating the sum of squares of the errors and using the constraints $\sum_i t_i = \sum_j b_j = 0$, the reduced normal equations for the treatment and regression coefficient estimates can be written as

$$(2.8)\quad \frac{r(k-1)}{k}\hat{t}_i - \sum_{s\neq i}\frac{\lambda_{is}}{k}\hat{t}_s + \sum_w \hat{B}_w Q_{iw} = Q_{i0} \quad (i = 1, 2, \cdots, v)$$

$$(2.9)\quad \sum_u E_{uw}\hat{B}_u = E_{0w} \qquad w = 1, 2, \cdots, p.$$

where $\lambda_{is} = \sum_j n_{ij}n_{sj}$ = number of blocks in which treatments i and

s appear together. Hence, the solution for the $\hat{t}_i$ can be alternatively written as

$$\hat{t}_i = \hat{t}_{i0} - \sum_w \hat{B}_w \hat{t}_{iw} \qquad (i = 1, 2, \cdots, v) \tag{2.10}$$

where $\hat{t}_{i0}$ is the treatment estimate if all $\hat{B}_w = 0$. The $\hat{t}_{iw}$ are a similar set of quantities obtained by replacing Q_{i0} by Q_{iw} for $w = 1, 2, \cdots, p$. Further, by some elementary algebra, we can demonstrate that $\hat{t}_{i0}$ and $\hat{B}_w$ are uncorrelated. Thus, the procedures for obtaining the estimates of the adjusted treatment effects, (adjusted or corrected for the effects of the concomitant variates) are: first, find the usual treatment estimates ($\hat{t}_{i0}$) neglecting the presence of the concomitant variates; second, repeat the calculations with the same formulae, but using the concomitant variates $x_{ij}^{(w)}$ ($w = 1, 2, \cdots, p$) in place of y_{ij}, and thus obtain the quantities $\hat{t}_{iw}$. Substitution in equation (2.10) gives the final result.

The corresponding residual sum of squares after adjusting for the effects of the concomitant variates is

$$S_e = E_{00} - \sum_w \hat{B}_w E_{0w} \tag{2.11}$$

having $n_e = (bk - b - v - p + 1)$ degrees of freedom. If the concomitant variates had been completely neglected, then the residual variance would be E_{00} which would be an inflated estimate of the residual sum of squares.

In the analysis of variance it is common to test the null hypothesis that all treatment effects are the same, e.g. $H_0(t_1 = t_2 = \cdots = t_v = 0)$. This can be tested by the F-ratio having $(v - 1)$ and n_e degrees of freedom,

$$F = \frac{S_t/(v - 1)}{S_e/n_e} \tag{2.12}$$

where the treatment sum of squares, S_t, is

$$S_t = T_{00} + \sum_w \hat{B}_w E_{0w} - \sum_w \tilde{B}_w [E_{0w} + T_{0w}] \tag{2.13}$$

and $\tilde{B}_w$ satisfies the linear equations

$$\sum_w \tilde{B}_w [E_{uw} + T_{uw}] = E_{0u} + T_{0u} \qquad (u = 1, 2, \cdots, p). \tag{2.14}$$

A test for the null hypothesis that the concomitant variates have no effect, i.e. $H_0(B_1 = B_2 = \cdots = B_p = 0)$ can be made using the

F-ratio

$$F = \frac{(\sum \hat{B}_w E_{0w})/p}{S_e/n_e} \tag{2.15}$$

having p and n_e degrees of freedom.

The variance of the difference between two treatment estimates $\hat{t}_a$ and $\hat{t}_b$ (say) is given by

$$\text{var}\,(\hat{t}_a - \hat{t}_b) = \text{var}\,(\hat{t}_{a0} - \hat{t}_{b0}) + \sum_{u,w} (\hat{t}_{au} - \hat{t}_{bu})(\hat{t}_{aw} - \hat{t}_{bw}) c_{uw}\sigma^2 \tag{2.16}$$

where var $(\hat{t}_{a0} - \hat{t}_{b0})$ is the variance between the treatment estimates ignoring the concomitant information and c_{uw} are the elements of the inverse matrix to $|| E_{uw} ||$.

The variances and covariances for the $\hat{B}_w$ are

$$\begin{cases} \text{var}\,\hat{B}_u = c_{uu}\sigma^2 \\ \text{cov}\,(\hat{B}_u\,, \hat{B}_w) = c_{uw}\sigma^2. \end{cases} \tag{2.17}$$

The case most often met in practice is when there is only one concomitant variate, i.e. $p = 1$. The computational procedure for this case is illustrated by a numerical example in section 6.

3. INTER-BLOCK ANALYSIS

In the case of incomplete block designs (designs in which the number of treatments is larger than the number of experimental units in a block, i.e. $v > k$) additional information can be recovered from the block totals provided one can assume that the block effects are random variables. One can use this additional inter-block information in two ways; e.g. (1) to find inter-block estimates of the treatments and regression coefficients which are independent of the intra-block estimates or (2) to find combined estimates which depend on both the intra- and inter-block information. Section 3.2 gives the formulae for finding the inter-block estimates and also shows how it may be possible to duplicate all statistical tests made in the intra-block analysis. Section 3.3 gives the necessary formulae for obtaining the combined estimates. These are more precise than if either the intra- or inter-block estimates had been used alone.

With respect to the non-covariance situation, most statisticians agree that the inter-block analysis may be important if the number of blocks is "large" or if the variability between blocks is "small." However, in the analysis of covariance, the inter-block analysis will be important

if the variability of the concomitant variate is larger for "between blocks" as compared to the variability "within blocks." This situation may allow more precise inter-block estimates of the regression coefficients as compared to the corresponding intra-block estimates.

3.1 *Notation*

A block total (summing equation (2.1) over i) will have the underlying mathematical model

$$(3.1) \qquad Y_{.j} = k\mu + \sum_i n_{ij}t_i + k\sum_w B_w\bar{x}_{.j}^{(w)} + \sum_i n_{ij}\epsilon_{ij} + kb_j$$

where b_j and ϵ_{ij} are assumed to be mutually uncorrelated random variables such that

$$E(\epsilon_{ij}) = 0, \qquad \text{var}\,(\epsilon_{ij}) = \sigma^2$$
$$E(b_j) = 0, \qquad \text{var}\,(b_j) = \sigma_b^2$$

(For all tests of significance it is further assumed that ϵ_{ij} and b_j are normally distributed variates).

The model (3.1) assumes that the regression of $Y_{.j}$ on $X_{.j}^{(w)}$ is B_w which is the same regression as the regression of y_{ij} on $x_{ij}^{(w)}$ described in the intra-block analysis. Some statisticians, however, have advocated a more general model which allows the intra-block regression coefficients to be different from the inter-block regression. In this paper, all models are such that the intra-block regression is the same as that for the inter-block regression. It is difficult for this writer to visualize situations allowing separate regressions. That is, if the regression of y on x is B, it seems that the regression of their averages should also be the same. If the two independent (intra- and inter-block) estimates of the respective regression coefficients differ more than by their indicated precisions, this is evidence of the inapplicability of the model, rather than as evidence that there are separate intra- and inter-block regressions.

It will be convenient to denote by a prime (′) quantities derived for the inter-block analysis that have counter-parts in the intra-block analysis. Thus,

$$(3.2)\begin{cases} t_i' = \text{estimate for the } i\text{th treatment obtained only from inter-block analysis.} \\ B_w' = \text{estimate for } B_w \text{ obtained from the inter-block analysis.} \\ t_{i0}' = \text{estimate for the } i\text{th treatment obtained from the inter-block analysis if } B_w = 0 \text{ for } w = 1, 2, \cdots, p. \\ t_{iw}' = \text{quantity obtained by using } x_{ij}^{(w)} \text{ in place of } y_{ij} \text{ and solving for } t_{i0}'. \end{cases}$$

Also let

$$(3.3)\quad \begin{cases} Q'_{i0} = Y_{i.} - Q_{i0} - r\bar{y}_{..} = \sum_j n_{ij}\bar{y}_{.j} - r\bar{y}_{..} \\ Q'_{iw} = X_{i.}^{(w)} - Q_{iw} - r\bar{x}_{..}^{(w)} = \sum_j n_{ij}\bar{x}_{.j}^{(w)} - r\bar{x}_{..}^{(w)} \\ T'_{00} = \sum_i t'_{i0}Q'_{i0}\ , \quad T_{uw} = \sum_i t'_{iu}Q'_{iw}\ , \quad T'_{0w} = \sum_i t'_{i0}Q'_{iw} \\ E'_{00} = b_{00} - T'_{00}\ , \quad E'_{uw} = b_{uw} - T'_{uw}\ , \quad E'_{0w} = b_{0w} - T'_{0w} \end{cases}$$

Since both the intra- and inter-block analyses yield independent treatment estimates, these estimates can be combined to form a single estimate having a smaller variance than either the intra- or inter-block estimates. A bar ( ) is used to denote those quantities which occur in the combined estimate. Thus

$$(3.4)\quad \begin{cases} \bar{t}_i = \text{combined estimate for treatment } i. \\ \bar{t}_{iw} = \text{combined treatment estimate for treatment } i \text{ if } B_w = 0 \\ \qquad \text{for } (i = 1, 2, \cdots, p). \\ \bar{B}_w = \text{combined estimate for } B_w\ . \end{cases}$$

Also let

$$(3.5)\quad \begin{cases} W = \dfrac{1}{\sigma^2} \\ W' = \dfrac{1}{\sigma^2 + k\sigma_b^2} \end{cases}$$

and define

$$(3.6)\quad \begin{cases} \bar{Q}_{i0} = WQ_{i0} + W'Q'_{i0} \\ \bar{Q}_{iw} = WQ_{iw} + W'Q'_{iw} \end{cases}$$

$$(3.7)\quad \begin{cases} \bar{E}_{00} = WE_{00} + W'E'_{00} \\ \bar{E}_{0w} = WE_{0w} + W'E'_{0w} \\ \bar{E}_{uw} = WE_{uw} + W'E'_{uw}\ . \end{cases}$$

3.2 *Estimates based only on inter-block information*

The estimates are obtained by minimizing the residual sum of squares $(1/k)\sum_i [kb_i + \sum_i n_{ij}\epsilon_{ij}]^2$ with respect to each of the unknown parameters. This results in the normal equations

$$(3.8)\quad \frac{r}{k}t'_i + \sum_{s \neq i} \frac{\lambda_{is}}{k} t'_s + \sum_w B'_w Q'_{iw} = Q'_{i0} \qquad i = 1, 2, \cdots, v$$

$$(3.9)\quad \sum_u B'_u E'_{uw} = E'_{0w} \qquad w = 1, 2, \cdots, p.$$

from which the t'_i can be written as

$$t'_i = t'_{i0} - \sum_w B'_w t'_{iw} \quad . \tag{3.10}$$

The residual sum of squares will have $n'_e = (b - v - p)$ degrees of freedom and is given by

$$S'_e = b_{00} - T'_{00} - \sum_w B'_w T'_{0w} \quad . \tag{3.11}$$

S'_e will be independent of S_e, has the expected value $E(S'_e) = n'_e(\sigma^2 + k\sigma_b^2)$, and follows a $(\sigma^2 + k\sigma_b^2)\chi^2$ distribution with n'_e degrees of freedom.

Thus all results derived for the intra-block analysis can be found for the inter-block analysis. Furthermore, the parameter estimates t'_i , B'_w are independent of the corresponding intra-block estimates. The variance of the difference between two treatment estimates t'_a and t'_b (say) is

$$\begin{aligned} \text{var}\,(t'_a - t'_b) &= \text{var}\,(t'_{a0} - t'_{b0}) \\ &\quad + \sum_{u,w} (t'_{aw} - t'_{bw})(t'_{au} - t'_{bu})c'_{uw}(\sigma^2 + k\sigma_b^2) \end{aligned} \tag{3.12}$$

where c'_{uw} are the elements of the inverse matrix to $|| E'_{uw} ||$.

Furthermore all tests of significance made in the intra-block analysis can be duplicated using only the inter-block estimates. Thus the equations (2.12)-(2.17) carry over directly by substituting primed (′) quantities for circumflex (^) quantities, $n'_e = b - p - v$ for n_e , and $(\sigma^2 + k\sigma_b^2)$ for σ^2. In addition, all of these inter-block tests of significance are independent of their intra-block counterparts.

Thus, from the b block totals we can estimate $(p + v)$ parameters, i.e., the grand mean μ, the parameters t_i $(i = 1, 2, \cdots, v$, of which $v - 1$ are independent) and the regression coefficients B_w $(w = 1, 2, \cdots, p)$. Hence, the number of blocks must be equal to or greater than the total number of parameters, i.e. $b \geq (v + p)$. Furthermore, if the blocks can be arranged in m replication groups, these $(m - 1)$ degrees of freedom can only be used for estimating the regression coefficients B_w and the variability between blocks. Therefore, a necessary and sufficient condition for estimating all $(p + v)$ parameters is that

$$b - v \geq \text{maximum}\,(m - 1, p).$$

In order to obtain an estimate of $(\sigma^2 + k\sigma_b^2)$ and thus be able to duplicate all tests in the inter-block analysis, it is necessary that the inequality sign hold or that $m > 1$. Further details can be found in [22].

3.3 *The combination of significance tests*

Since there exist two independent variance-ratio tests for testing any linear hypothesis concerning the treatments or regression coefficients,

it seems desirable to combine the two tests into a single test. A proposed procedure is to let P_i $(i = 1, 2)$ refer to the probabilities that the calculated F_i value will be exceeded under the null hypothesis, i.e.

$$P_i = P\{F \geq F_i \mid H_0\} \qquad i = 1, 2 \tag{3.13}$$

where the subscripts 1 and 2 refer to the intra- and inter-block analyses respectively. Then a combined test can be made using the statistic

$$C = P_1 P_2^{\theta} \tag{3.14}$$

where

$$\theta = \frac{1 - E}{E} \frac{\sigma^2}{\sigma^2 + k\sigma_b^2}, \tag{3.15}$$

and E refers to the efficiency of the experiment design. (The null hypothesis is rejected with level of significance α if $C < C_\alpha$). In most situations $\sigma^2/(\sigma^2 + k\sigma_b^2)$ will be replaced by its estimate $s_e^2/s_{e'}^2$. Then the test will only be an approximate one. The power of this test is investigated in [23]. A table of critical values of C_α for $\alpha = .01, .05$ is given in [22]. Other methods for combining significance tests are discussed by Birnbaum [4], Mosteller and Bush [16], and Wallis [20].

3.4 *Combined estimates*

The most precise treatment estimates are obtained by weighting both the intra- and inter-block estimates by the reciprocal of their respective variances. This results in estimates of the form

$$\hat{l}_i = \hat{l}_{i0} - \sum_u \bar{B}_w \hat{l}_{iw} \tag{3.16}$$

where the $\bar{B}_w$ satisfy the equations

$$\sum_u \bar{B}_u \bar{E}_{uw} = \bar{E}_{0w} \qquad (w = 1, 2, \cdots, p). \tag{3.17}$$

In the practical situation the weights

$$W = 1/\sigma^2, \quad W' = \frac{1}{\sigma^2 + k\sigma_b^2}$$

are not known and have to be estimated. For this purpose we have estimates of σ^2 and $\sigma^2 + k\sigma_b^2$ which follow independent $\sigma^2\chi^2$ and $(\sigma^2 + k\sigma_b^2)\chi^2$ distributions respectively when $b > \max(p, m - 1)$ or when $m > 1$. If the inter-block analysis is made without this condition holding, then no estimate of $(\sigma^2 + k\sigma_b^2)$ independent of S_e is available. Then one can determine an estimate of $(\sigma^2 + k\sigma_b^2)$, which is not independent of S_e, by an easily apparent modification of the usual analysis

of variance formulae applied to the recovery of inter-block information, cf. Federer [11].

4. FORMULAE FOR APPLICATION TO BALANCED AND PARTIALLY BALANCED INCOMPLETE BLOCK DESIGNS

For convenience, the quantities $\hat{t}_{i0}$, t'_{i0} , $\bar{t}_{i0}$ the variance of a difference, and the efficiency factor E for the balanced and partially balanced designs (two associate classes) are listed below

Balanced incomplete block designs

$$\hat{t}_{i0} = \frac{Q_{i0}}{Er} \tag{4.1}$$

$$t'_{i0} = \frac{Q'_{i0}}{(1 - E)r} \tag{4.2}$$

$$\bar{t} = \frac{\bar{Q}_{i0}}{r[EW + (1 - E)W']} \tag{4.3}$$

$$E = \frac{v(k - 1)}{k(v - 1)} \tag{4.4}$$

$$\text{var}\,(\hat{t}_{a0} - \hat{t}_{b0}) = \frac{2\sigma^2}{Er} \tag{4.5}$$

$$\text{var}\,(t'_{a0} - t'_{b0}) = \frac{2(\sigma^2 + k\sigma_b^2)}{(1 - E)r} \tag{4.6}$$

Partially balanced incomplete block designs

$$\hat{t}_{i0} = \frac{1}{r(k - 1)}\,[kQ_{i0} + c_1S_1(Q_{i0}) + c_2S_2(Q_{i0})], \tag{4.7}$$

$$t'_{i0} = \frac{1}{r}\,[kQ'_{i0} + c'_1S_1(Q'_{i0}) + c'_2S_2(Q'_{i0})], \tag{4.8}$$

$$\bar{t}_{i0} = \frac{1}{r[W(k - 1) + W']}\,[k\bar{Q}_{i0} + d_1S_1(\bar{Q}_{i0}) + d_2S_2(\bar{Q}_{i0})], \tag{4.9}$$

$$E = \frac{(v - 1)(k - 1)}{n_1(k - c_1) + n_2(k - c_2)} \tag{4.10}$$

where $S_m(Q_{i0})$, $S_m(Q'_{i0})$, $S_m(\bar{Q}_{i0})$ is the sum of the mth associates Q_{i0} , Q'_{i0} , $\bar{Q}_{i0}$ for treatment i_i , c'_1 , c'_2 are defined by

$$c'_m = \frac{c_m\Delta - r\lambda_m}{\Delta - rH + r^2} \qquad (m = 1, 2) \tag{4.11}$$

and the quantities c_m , d_m , n_m , T_m , Δ, and H are the usual parameters

for partially balanced designs, cf. Bose and Shimamoto [5] or Bose, Clatworthy, and Shrikhande [6].

If a and b are mth associates, then

$$\text{var}\,(\hat{t}_{a0} - \hat{t}_{b0}) = \frac{2\sigma^2(k - c_m)}{r(k-1)} \tag{4.12}$$

$$\text{var}\,(t'_{a0} - t'_{b0}) = \frac{2(\sigma^2 + k\sigma_b^2)(k - c'_m)}{r} \tag{4.13}$$

For some partially balanced designs, it will not be possible to estimate all $v - 1$ treatment contrasts. This will occur when the denominator of (4.11) is zero, i.e. $\Delta - rH + r^2 = 0$. When this condition holds, it implies that there exist (say) s orthogonal contrasts among the block totals, having expectations independent of the treatments and $s > b - v$. These s contrasts can be used for estimating the p regression coefficients (provided $s \geq p$), and $\sigma^2 + k\sigma_b^2$. The remaining $(b - s) - 1$ contrasts among the blocks will have expectations dependent on the treatments. However, the number of orthogonal treatment contrasts is $v - 1$ which is greater than the $(b - s) - 1$ contrasts among the block totals available for estimating all the treatment contrasts, i.e. $v > b - s$. Therefore not all treatment contrasts will be estimable. This will be especially true for the singular and semi-regular group divisible designs.

5. ANALYSIS OF COVARIANCE WHEN THE REGRESSION COEFFICIENTS DEPEND ON THE BLOCKS

All the techniques for the analysis of covariance discussed in previous sections were based on the assumption that the regression coefficients are independent of the blocks. In some experimental situations this assumption is not satisfied. This section outlines the appropriate covariance analysis for *one* concomitant variate when the regression coefficient is dependent on the block.

This model coincides with that suggested by C. P. Cox [10] where the mathematical model (excluding the t_i) can be regarded as a curved sheet representing the background of the experimental material in three dimensions. If the concomitant variate represents rows and the blocks represent columns, the indicated model would allow the columns to be discrete and the variation within them represented by linear regressions.

Let the mathematical model underlying the measurement for the ith treatment in the jth block be

$$y_{ij} = \mu + t_i + b_j + B_j(x_{ij} - \bar{x}_{.j}) + \epsilon_{ij} \tag{5.1}$$

where the quantities μ, t_i, and b_j have the same meaning as in previous sections, but $B_j (j = 1, 2, \cdots, b)$ is a regression coefficient *associated with the jth block*.

The normal equations are obtained by differentiating the residual sum of squares

$$S = \sum_{i,j} n_{ij}[(y_{ij} - \mu - t_i - b_j - B_j(x_{ij} - \bar{x}_{.j})]^2$$

with respect to μ, t_i, b_j, and B_j. Using the restraints $\sum_i t_i = \sum_j b_j = 0$, the normal equations take the form

$$\bar{y}_{..} = \hat{\mu} \tag{5.2}$$

$$Y_{i.} = r(\hat{\mu} + \hat{t}_i) + \sum_j n_{ij}\hat{b}_j + \sum_j n_{ij}\hat{B}_j(x_{ij} - \bar{x}_{.j}) \tag{5.3}$$

$$Y_{.j} = k(\hat{\mu} + \hat{b}_j) + \sum_i n_{ij}\hat{t}_i \tag{5.4}$$

$$\sum_i n_{ij}y_{ij}(x_{ij} - \bar{x}_{.j}) = \sum_i n_{ij}(x_{ij} - \bar{x}_{.j})\hat{t}_i + \hat{B}_j \sum_i n_{ij}(x_{ij} - \bar{x}_{.j})^2 \tag{5.5}$$

Substituting (5.2) and (5.4) in (5.3) one can write (5.3) as

$$\left(\frac{r(k-1)}{k}\right)\hat{t}_i - \left(\frac{1}{k}\right)\sum_{s \neq i} \lambda_{is}\hat{t}_s + \sum_j \hat{B}_j q_{ij} = Q_{i0} \qquad (i = 1, 2, \cdots, v) \tag{5.6}$$

where $q_{ij} = n_{ij}(x_{ij} - \bar{x}_{.j})$.

Define the quantities $\hat{\tau}_{ij}$ by

$$\left(\frac{r(k-1)}{k}\right)\hat{\tau}_{ij} - \left(\frac{1}{k}\right)\sum_{s \neq i} \lambda_{si}\hat{\tau}_{sj} = q_{ij} \tag{5.7}$$

$$(i = 1, 2, \cdots, v; j = 1, 2, \cdots, b).$$

Note that the $\hat{\tau}_{ij}$ are the quantities calculated from (5.6) by replacing Q_{i0} by q_{ij} and setting $B_j = 0$ for all j.

Then by a rearrangement the $\hat{t}_i$ can be written as

$$\hat{t}_i = \hat{t}_{i0} - \sum_j \hat{B}_j\tau_{ij} \tag{5.8}$$

where $\hat{t}_{i0}$ are as defined in section 2. Substituting (5.8) in (5.5) results in the equations for $\hat{B}_j$ written in the simple form

$$\hat{B}_j E_{jj} - \sum_{h \neq j} \hat{B}_h T_{hj} = E_{jy} \qquad (j = 1, 2, \cdots, b) \tag{5.9}$$

where

$$(5.10)\qquad \begin{cases} T_{hj} = \sum_i \hat{\tau}_{ij} q_{ih} = \sum_i \hat{\tau}_{ih} q_{ij} \\ E_{ij} = \sum_i q_{ij}^2 - T_{ij} \\ E_{iy} = \sum_i q_{ij} y_{ij} - \sum_i \hat{t}_{i0} q_{ij} \ . \end{cases}$$

The variance of the difference between two treatment estimates is

$$(5.11)\qquad \operatorname{var}(\hat{t}_a - \hat{t}_b) = \operatorname{var}(\hat{t}_{ay} - \hat{t}_{by}) + \sum_{h,j} (\hat{\tau}_{ah} - \hat{\tau}_{bh})(\hat{\tau}_{bj} - \hat{\tau}_{aj}) c_{hj} \sigma^2$$

where c_{hj} are the elements of the inverse matrix arising from the set of linear equations given by (5.9).

The computations can conveniently be arranged in an analysis of variance table as shown in Table 1.

The test of significance for the null hypothesis that all treatment effects are zero, i.e. $H_0(t_1 = t_2 = \cdots = t_v = 0)$ is given by the F-ratio $F = s_t^2 / s_e^2$ having $(v - 1)$ and $n_e = (bk - 2b - v + 1)$ degrees of freedom.

A test for the null hypothesis that the regression coefficients do not depend on the blocks is given by the F-ratio $F = s_r^2/s_e^2$ having $(b - 1)$ and n_e degrees of freedom.

6. EXAMPLE WITH ONE CONCOMITANT VARIATE

The usual application of the analysis of covariance, as first described by Fisher, is to adjust the final estimates by means of a concomitant set of data in order to improve the precision of the final estimates. However, another use for the analysis of covariance (or analysis of covariance type calculations) is to employ a general regression model to relate a dependent variable in terms of both qualitative and quantative variables. This section contains such an example.

A certain film-type composition resistor used in electronic equipment is of the type which is mounted on a small ceramic plate. An investigation was recently conducted at the National Bureau of Standards to determine the effects of four different geometrical shapes on the current noise of these resistors. These geometric shapes were taken to be rectangular parallelepipeds (all having the same thickness) formed by taking all four combinations of 2 widths (w_1 , w_2) and 2 lengths (l_1 , l_2). For convenience the four combinations of length and width, $l_1 w_1$, $l_1 w_2$, $l_2 w_1$, $l_2 w_2$, will be denoted by a, b, c, and d respectively.

Since for this type of resistor the noise of an individual resistor varies with its resistance, an appropriate mathematical model underlying the measurements is

$$y_{ij} = \mu + t_i + b_j + B x_{ij} + \epsilon_{ij}$$

TABLE 1

ANALYSIS OF COVARIANCE WITH REGRESSION COEFFICIENTS DEPENDENT ON BLOCK

Source	Degrees of freedom	Sum of squares	Mean square
Treatments (adjusted for blocks and regression cofficients)	$(v-1)$	$S_t = T_{vv} + \sum_i \hat{B}_i E_{iv} - \sum_i \frac{[E_{iv}+T_{iv}]^2}{E_{ii}+T_{ii}}$	$s_t^2 = \frac{S_t}{(v-1)}$
Blocks (unadjusted)	$(b-1)$	$b_{vv} = \sum_i \frac{Y_{\cdot i}^2}{k} - \frac{Y_{\cdot\cdot}^2}{bk}$	
Differential regression coefficients	$(b-1)$	$S_r = \sum_i \hat{B}_i E_{iv} - \frac{E_{xv}^2}{E_{xx}}$	$s_r^2 = \frac{S_r}{b-1}$
Mean regression	1	$\frac{E_{xv}^2}{E_{xx}}$	$\frac{E_{xv}^2}{E_{xx}}$
Error	$n_e = bk - 2b - v + 1$	$S_e = E_{vv} - \sum_i \hat{B}_i E_{iv}$	$s_e^2 = \frac{S_e}{n_e}$

where y_{ij} is the logarithm of the noise measurement of the ith type resistor (treatment) in the jth plate (block), x_{ij} is the logarithm of the resistance for this resistor, μ is a constant common to all observations, t_i is the effect of the ith treatment, and B is a constant of proportionality between "noise" and "resistance". Furthermore, the plate effects $\{b_j\}$ and the $\{\epsilon_{ij}\}$ are assumed to be sequences of independent random variables following normal distributions with mean zero and variance σ_b^2 and σ^2 respectively.

This experimental situation does not conform to the usual assumptions underlying a covariance analysis, as here, the concomitant variate (resistance) depends partly on the shape of the resistor. However, from previous work it is known that a linear relation exists between noise and resistance (both on logarithm scales). The purpose of this experiment was to determine if the intercept of this relation varies with the geometry of the resistor. This is equivalent to testing whether all treatment effects t_i are equal to zero. Thus, the fact that the value of the resistance depends partly on the shape of the resistor in no way invalidates the test that the intercept $(\mu + t_i)$ varies with the shape of the resistor.

Since only three different resistors can be mounted on any one ceramic plate, a balanced incomplete block design involving 36 resistors mounted on 12 plates was used for the basic experimental plan. The plates determine a "natural" grouping of the experimental material and hence can be regarded as constituting the "blocks" in the experiment. The parameters of the experimental design are

$$v = 4, \qquad k = 3, \qquad b = 12, \qquad r = 9, \qquad \lambda = 6, \qquad E = \frac{8}{9}.$$

Table 2A summarizes the coded measurements and gives the experimental plan. The four treatments are denoted by the letters a, b, c, d and the numbers in parentheses represent the values for the concomitant variate.

The values for Q_{iz}, Q'_{iz}, $\hat{t}_{iz}$ for $z = x, y$ are summarized in Table 2B where the treatment estimates, ignoring the concomitant variate, are obtained using equations (4.1) and (4.2) e.g.

$$(6.1) \qquad \begin{cases} \hat{t}_{iz} = \dfrac{Q_{iz}}{8} & \text{(intra-block)} \\ t'_{iz} = Q'_{iz} & \text{(inter-block)} \end{cases}$$

TABLE 2A

EXPERIMENTAL PLAN

Blocks	Observations and experimental plan					
		1.11		.95		.82
1	a	(.98)	c	(.66)	d	(.47)
		1.22		.97		1.70
2	b	(.88)	d	(.58)	a	(1.14)
		1.52		1.60		1.11
3	c	(.79)	a	(1.08)	b	(.84)
		1.18		1.22		1.54
4	d	(.58)	b	(.88)	c	(.80)
		1.57		1.33		.93
5	a	(1.08)	c	(.79)	d	(.49)
		1.12		.93		1.69
6	b	(.82)	d	(.51)	a	(1.11)
		1.26		1.35		1.06
7	c	(.69)	a	(1.00)	b	(.80)
		1.12		1.09		1.29
8	d	(.48)	b	(.79)	c	(.70)
		1.60		1.82		1.09
9	a	(1.12)	c	(.82)	d	(.57)
		1.03		.94		1.66
10	b	(.80)	d	(.52)	a	(1.10)
		1.24		1.31		1.11
11	c	(.71)	a	(1.00)	b	(.82)
		.93		1.09		1.27
12	d	(.47)	b	(.78)	c	(.68)

TABLE 2B

TREATMENT ESTIMATES IGNORING CONCOMITANT VARIATE

Treatment	Intra-block				Inter-block	
	Q_{iy}	$\hat{t}_{iy}$	Q_{ix}	$\hat{t}_{ix}$	t'_{iy}	t'_{ix}
a	2.243 333	.280 417	2.220 000	.277 500	.154 167	.307 500
b	−1.133 333	−.141 667	0.293 333	.036 667	−.009 167	.034 167
c	1.050 000	.131 250	−0.316 667	−.039 583	−.022 500	−.125 833
d	−2.160 000	−.270 000	−2.196 667	−.274 583	−.122 500	−.215 833

The computational procedure for the intra-block analysis of variance proceeds in a similar manner as for the simple randomized block case. The appropriate analysis of variance of the sums of products and cross

products can be summarized as in Table 2C. The treatment mean square for testing the null hypothesis that all treatment effects are the same, is most conveniently obtained by forming a reduced analysis of variance which is summarized in Table 2D.

The analogous results for the inter-block analyses can conveniently be arranged in a single table. These results are summarized in Table 2E.

TABLE 2C

AUXILIARY INTRA-BLOCK ANALYSIS OF VARIANCE

Source of Variation	Degrees of freedom	Sum of squares and products y^2	xy	x^2
Blocks (unadjusted for treatment)	11	0.641 898	0.258 630	0.222 631
Treatments (adjusted for blocks)	3	1.510 637	1.132 506	1.242 508
Error (intra-block)	21	0.278 896	0.038 661	0.011 025
Total	35	2.431 431	1.429 797	1.476 164

The reduced error sum of squares, corrected for the effects of the concomitant variate, in Tables 2D and 2E are obtained from the formulae

$$(6.2)\qquad \begin{cases} \text{(intra-block)} \quad S_e = E_{yy} - \dfrac{E_{xy}^2}{E_{xx}} = 0.143\,325 \quad \text{d.f.} = 20 \\ \text{(inter-block)} \quad S_e' = E_{yy}' - \dfrac{E_{xy}'^2}{E_{xx}'} = 0.087\,402 \quad \text{d.f.} = 7 \end{cases}$$

Thus the estimates of B, for the intra- and inter-block analyses respectively, result in

$$(6.3)\qquad \begin{cases} \hat{B} = \dfrac{E_{xy}}{E_{xx}} = 3.506\,667 \\ B' = \dfrac{E_{xy}'}{E_{xx}'} = 2.826\,268 \end{cases}$$

Hence, using (6.3) the treatment estimates corrected for the concomitant variate can be calculated. The second and third columns of Table 2F summarize the corrected treatment estimates for both the intra- and inter-block analyses. The fourth column gives the combined treatment estimates (equation 3.16) where the weights W and W' are taken to be

TABLE 2D

INTRA-BLOCK REDUCED ANALYSIS OF VARIANCE

Source	Degrees of freedom	Sum of squares and cross products y^2	xy	x^2	Degrees of freedom	Reduced sum of squares	Mean square
Treatments (adjusted)	3	1.510 637	1.132 506	1.242 508	3	0.551 995	0.183 998
Error (intra-block	21	0.278 896	0.038 661	0.011 025	20	0.143 325	0.007 166
(Treatments + error)		1.789 553	1.171 167	1.253 533		0.695 320	

TABLE 2E

INTER-BLOCK ANALYSIS OF VARIANCE

Source of variation	Degrees of freedom	Sum of squares and products y^2	xy	x^2	Degrees of freedom	Reduced sum of squares	Mean square
Treatments (adjusted)	3	.039 364	.076 364	.158 141	3	0.254 046	0.084 682
Error (inter-block)	8	.602 534	.182 266	.064 490	7	0.087 402	0.012 486
Blocks (unadjusted)	11	.641 898	.258 630	.222 631	10	0.341 448	

$$(6.4)\qquad \begin{cases} w = \dfrac{1}{s_e^2} = 139.542\,997 \\ w' = \dfrac{1}{s_e'^2} = 80.089\,700 \end{cases}$$

TABLE 2F

TREATMENT ESTIMATES CORRECTED FOR THE CONCOMITANT VARIATE

Treatments	$\hat{t}_i$ (intra-block)	t_i' (inter-block)	$\bar{t}_i$ (combined)	$\hat{\tau}_i$
a	−0.692 683	−0.714 910	−0.561 646	−0.503 872
b	−0.270 246	−0.105 732	−0.241 653	−0.245 296
c	0.270 054	0.333 138	0.256 232	0.243 122
d	0.692 871	0.487 502	0.547 068	0.506 045

The estimated variances for $\hat{B}$ and B' are

$$(6.5)\qquad \begin{cases} \text{var}\,\hat{B} = \dfrac{s_e^2}{E_{xx}} = \dfrac{.007\,166}{.011\,025} = 0.649\,977 \\ \text{var}\,B' = \dfrac{s_e'^2}{E_{xx}'} = \dfrac{.012\,486}{.064\,490} = 0.193\,611 \end{cases}$$

Note that the intra-block estimated variance is over three times larger than the inter-block estimated variance.

Since the variance between two adjusted treatment estimates is different for each comparison, Finney [12] has suggested using an average value of $(\hat{t}_a - \hat{t}_b)^2$. Hence for the intra-block analysis we have

$$(6.6)\qquad \text{"average" var}\,(\hat{t}_a - \hat{t}_b) = \frac{2\sigma^2}{Er}\left[1 + \frac{T_{xx}}{(v-1)E_{xx}}\right] = 0.069\,111.$$

A similar formula holds for the inter-block analysis

$$(6.7)\qquad \text{"average" var}\,(t_a' - t_b') = \frac{2(\sigma^2 + k\sigma_b^2)}{(1-E)r}\left[1 + \frac{T_{xx}'}{(v-1)E_{xx}'}\right] = 0.045\,384.$$

However, if the treatments adjusted for the concomitant variate are estimated using

$$(6.8)\qquad \hat{\tau}_i = \hat{t}_{iy} - B'\hat{t}_{ix}$$

where B' is the inter-block estimate of B, then the average variance of a

comparison between two treatments is

$$\text{(6.9)} \qquad \text{var}\,(\hat{\tau}_a - \hat{\tau}_b) = \frac{2}{Er}\left[\sigma^2 + \frac{(\sigma^2 + k\sigma_b^2)T_{rr}}{E'_{xx}(v-1)}\right] = 0.017\,926.$$

Thus, the estimated variance is reduced by almost 3/4 over that of the usual intra-block variance. The last column of Table 2F gives the treatment estimates using (6.8).

The two independent tests of significance for the null hypothesis that all treatment effects are the same are obtained by testing the error mean square against the treatment mean square obtained from Tables 2D and 2E, e.g.

$$\text{(6.10)} \qquad \begin{cases} \text{(intra-block)} \quad F = \dfrac{0.183\,998}{0.007\,166} = 25.68 \quad \text{d.f.} = 3, 20 \\ \text{(inter-block)} \quad F = \dfrac{0.084\,682}{0.012\,486} = 6.78 \quad \text{d.f.} = 3, 7 \end{cases}$$

Both F-ratios in (6.10) are highly significant. Thus, there is no real need to combine the two independent tests into a single test.

Instead the combination of the two independent tests will be illustrated by testing the interaction between lengths and widths. The contrast

$$\pi = (c - d) - (a - b)$$

can be used to test if an interaction exists between lengths and widths. If π is equal to zero, then there is no interaction between lengths and widths. The variance of the two independent estimates of π obtained from the intra- and inter-block analysis are

$$\text{(6.11)} \qquad \begin{cases} \text{var}\,\hat{\pi} = \dfrac{4\sigma^2}{Er} + \dfrac{[(\hat{t}_{ax} - \hat{t}_{bx}) - (\hat{t}_{cx} - \hat{t}_{dx})]^2}{E_{xx}}\,\sigma^2 \\ \text{var}\,\pi' = \dfrac{4(\sigma^2 + k\sigma_b^2)}{(1-E)r} + \dfrac{[(t'_{ax} - t'_{bx}) - (t'_{cx} - t'_{dx})]^2}{E'_{xx}}\,(\sigma^2 + k\sigma_b^2) \end{cases}$$

respectively. Thus two independent t-tests can be made for the null hypothesis that $\pi = 0$ against the alternative hypothesis that $\pi \neq 0$.

When combining tests of significance in cases where the alternative hypothesis is two sided, a preferred procedure (cf. Fisher [14]) is to choose a direction as positive, obtain the associated probabilities of getting the observed value or a larger one, and combine the resulting probabilities. After combining these probabilities, a two-tailed test can be used for the final result. This technique will be illustrated here.

The two estimates for π and their associated estimated variances are

$$\hat{\pi} = -0.000\,380, \qquad \pi' = +0.454\,814,$$
$$\operatorname{var}\hat{\pi} = 0.003\,605, \qquad \operatorname{var}\pi' = 0.056\,451,$$

which give the t-ratios

$$\begin{cases} \text{(intra-block)} & t_1 = -0.0063 \quad \text{d.f.} = 20 \\ \text{(inter-block)} & t_2 = 1.9143 \quad \text{d.f.} = 7. \end{cases}$$

Hence, the probabilities of obtaining t-ratios equal to or larger than those above are

$$\begin{cases} P_1 = 0.50238 \\ P_2 = 0.04857. \end{cases}$$

The value of the weight function θ is

$$\theta = \frac{1 - E}{E} \frac{s_e^2}{s_e'^2} = .06$$

and thus

$$C = P_1 P_2^{\theta} = 0.445.$$

Since the alternative hypothesis is actually two-sided, the critical values in [23] refer to the .02 and .10 levels of significance. If we adopt (say) a level of significance of $\alpha = .10$, the critical value corresponding to $\theta = .06$ is $C_{.10} = .047$ (obtained by linear interpolation). Thus, one would accept the null hypothesis that there is no interaction between lengths and widths.

ACKNOWLEDGEMENT

I would like to thank Mr. George Conrad for the use of the data in the illustrative example. Also, I would like to thank Professor Oscar Kempthorne and Dr. Allyn Kimball for contributing many constructive suggestions which served to improve the paper. While I did not always agree with their comments, they always served to sharpen my thinking.

REFERENCES

[1] Anderson, R. L. and Bancroft, T. A. [1952] *Statistical theory in research*, McGraw-Hill Book Company, Inc., New York.

[2] Bartlett, M. S. [1936] A note on the analysis of covariance, *J. Agri. Sc., 26:* 488–491

[3] Bartlett, M. S. [1937] Some examples of statistical methods of research in agriculture and applied biology, *J. Roy. Stat. Soc. Supp., 4:* 137–169.

[4] Birnbaum, A. [1954] Combining independent tests of significance, *J. A. S. A.*, *49:* 559–574.

[5] Bose, R. C. and Shimamoto [1952] Classification and analysis of partially balanced incomplete block designs with two associate classes, *J. A. S. A.*, *47:* 151–184.

[6] Bose, R. C., Clatworthy, W. H. and Shrikhande, S. S. [1954] *Tables of partially balanced designs with two associate classes,* Institute of Statistics of the University of North Carolina, Reprint Series No. 50.

[7] Cochran, W. G. [1946] Analysis of covariance, *Institute of Statistics,* Mimeographed Series 6, University of North Carolina.

[8] Cochran, W. G., Cox, G. M., and Eckhardt, R. C. [1940] The Analysis of lattice and triple lattice experiments in corn varietal tests, *Iowa Agri. Expt. Stat. Res. Bull.*, 281.

[9] Cornish, E. A. [1940] The analysis of covariance in quasi-factorial designs, *Annals of Eugenics, 10:* 269–279.

[10] Cox, C. P. [1956] Latin square designs with individual gradients in one direction, *Nature, 177:* 1092.

[11] Federer, W. T. [1955] *Experimental design,* Macmillan Company, New York.

[12] Finney, D. J. [1946] Standard errors of yield adjusted for regression on an independent measurement, *Biometrics, 2:* 53–55.

[13] Fisher, R. A. [1946] *Statistical methods for research workers,* Oliver and Boyd Ltd., Edinburgh, 10th Edition.

[14] Fisher, R. A. [1948] Q 14, Questions and answers, *The American Statistician,* October: 30.

[15] Kempthorne, O. [1952] *The design and analysis of experiments,* John Wiley and Sons, Inc., New York.

[16] Mosteller, F., and Bush, R. R. [1954] Selected quantitative techniques, *Handbook of social psychology,* Chapter 8, Addison-Wesley, Cambridge.

[17] Rao, C. R. [1952] *Advanced statistical methods in biometric research,* John Wiley and Sons, New York: 118–124.

[18] Snedecor, G. W. [1946] *Statistical methods,* Iowa State College Press, Ames, Iowa, Fourth edition.

[19] Truett, J. T., and Smith, H. F. [1956] Adjustment by covariance and consequent tests of significance in split-plot experiments, *Biometrics, 12:* 23–39.

[20] Wallis, W. A. [1942] Compounding probabilities from independent significance tests, *Econometrica, 10:* 229–248.

[21] Wishart, J. [1950] Field Trials II: The analysis of covariance, *Commonwealth Bureau of Plant Breeding and Genetics,* Technical Communication No. 15.

[22] Zelen, M. [1957] The analysis of incomplete block designs, *J. A. S. A.*, *52:* 204–217.

[23] Zelen, M., and Joel, L. S., The weighted compounding of two probabilities from independent significance tests, Submitted to a technical journal.

VARIANCE AND COVARIANCE ANALYSES FOR UNBALANCED CLASSIFICATIONS*

WALTER T. FEDERER
Cornell University, Ithaca, N.Y., U.S.A.

INTRODUCTION

Yates [20, 21], Snedecor [16], Snedecor and Cox [17], Nair [12], Cochran [3], Stevens [18], Henderson [10], and Rao [15] have presented analysis of variance procedures for unbalanced classifications while Day and Fisher [5], Wilks [19], Hazel [9], Rao [15], Henderson [11], Das [4], and Federer [6, 7, 8] have discussed covariance analysis for unbalanced classifications. Bartlett [1, 2], Quenouille [14], Federer [7] and Outhwaite and Rutherford [13], among others, have discussed the use of dummy covariates to remove the effect of disproportion in particular unbalanced classifications.

The purpose of this paper is to present additional results for variance and covariance analyses with unbalanced classifications. In particular, the analyses are grouped as follows:

Case I—Interaction absent;
Case II—Interaction present; the effects assumed to be fixed effects;
Case III—Interaction present; the interaction effects and at least one of the main effects of the factors represented in the interactions assumed to be random effects.

A Case I variance analysis is known as "the method of fitting constants" and a Case II variance analysis is known as "the weighted squares of means analysis" in the literature [16, 17, 20, 21]. A Case III analysis of variance has been discussed in the literature for balanced classifications [e.g., see 7, ch. VIII], but not for unbalanced classifications. A covariance analysis for one-way classifications has been discussed by Day and Fisher [5]. Case I covariance analyses for unbalanced classifications have been presented by Das [4], Federer [6, 8], and Hazel [9]. Federer [6, 8] presented the Case II and the Case III linear covariance analyses for an unbalanced two-way classification, and included

*Paper presented at the joint meeting of the Biometric Society, ENAR, the American Statistical Association, and the Institute of Mathematical Statistics, Detroit, Sept. 7, 1956.
Paper No. 340 Dept. of Plant Breeding and No. 35 Biometrics Unit, Dept. of Plant Breeding, Cornell University.

illustrative numerical examples. The additional results included in the present paper represent extensions of previous work [4, 6, 8] to unbalanced q-way classifications with several covariates.

CASE I

The linear model for a two-way classification (Table 1) with a covariate is:

$$Y_{ijh} = \mu + \tau_i + \rho_j^{\circ} + \epsilon_{ijh} + \beta(X_{ijh} - \bar{x}), \tag{I-1}$$

where μ = an effect common to all observations, τ_i = an effect common to the ith level of the first factor or classification, ρ_j = an effect common to the jth level of the second factor or classification, ϵ_{ijh} = a random error component, $\bar{x}$ = arithmetic mean of the X's, and β = the true regression of Y on X estimated from the residual line of the analysis of covariance (Table 2). With multiple regression the linear model becomes:

$$Y_{ijh} = \mu + \tau_i + \rho_j + \epsilon_{ijh} + \sum_{g=1}^{b} \beta_g(X_{gijh} - \bar{x}_g), \tag{I-2}$$

where β_g is the true partial regression of Y on the gth independent variate, X_g, estimated from the residual line of the analysis of covariance and $\bar{x}_g$ is the arithmetic mean of the X's in the gth group.

For the linear model in formula (I-2), the normal equations for $\hat{\mu}$, $\hat{\tau}_i$, $\hat{\rho}_j$, and $\hat{\beta}_g$, respectively, are:

$$\sum_{i=1}^{v} \sum_{j=1}^{r} \sum_{h=1}^{n_{ij}} Y_{ijh} = Y_{\ldots} = n_{\ldots}\hat{\mu} + \sum_{i=1}^{v} n_{i.}\hat{\tau}_i + \sum_{j=1}^{r} n_{.j}\hat{\rho}_j \; ; \tag{I-3}$$

$$\sum_{j=1}^{r} \sum_{h=1}^{n_{ij}} Y_{ijh} = Y_{i\ldots} = n_{i.}(\hat{\mu} + \hat{\tau}_i) + \sum_{j=1}^{r} n_{ij}\hat{\rho}_j + \sum_{g=1}^{b} \hat{\beta}_g(X_{gi\ldots} - n_{i.}\bar{x}_g); \tag{I-4}$$

$$\sum_{i=1}^{v} \sum_{h=1}^{n_{ij}} Y_{ijh} = n_{.j}(\hat{\mu} + \hat{\rho}_j) + \sum_{i=1}^{v} n_{ij}\hat{\tau}_i + \sum_{g=1}^{b} \hat{\beta}_g(X_{g.j.} - n_{.j}\bar{x}_g); \tag{I-5}$$

$$\sum_{i} \sum_{j} \sum_{h} Y_{ijh}(X_{gijh} - \bar{x}_g) = \sum_{i=1}^{v} \hat{\tau}_i(X_{gi\ldots} - n_{i.}\bar{x}_g) + \sum_{j=1}^{r} \hat{\rho}_j(X_{g.j.} - n_{.j}\bar{x}_g) + \sum_{i} \sum_{j} \sum_{h} (X_{gijh} - \bar{x}_g) \sum_{g=1}^{b} \hat{\beta}_g(X_{gijh} - \bar{x}_g). \tag{I-6}$$

TABLE 1

YIELD AND NUMBER OF OBSERVATIONS FOR THE VARIATE Y AND FOR THE COVARIATE X FROM A TWO-WAY CLASSIFICATION (TOTALS PER n_{ij} OBSERVATIONS).

First category	Second category						Total
	1	2	...	j	...	r	
1	$n_{11}\ Y_{11.}\ X_{11.}$	$n_{12}\ Y_{12.}\ X_{12.}$		$n_{1j}\ Y_{1j.}\ X_{1j.}$		$n_{1r}\ Y_{1r.}\ X_{1r.}$	$n_{1.}\ Y_{1..}\ X_{1..}$
2	$n_{21}\ Y_{21.}\ X_{21.}$	$n_{22}\ Y_{22.}\ X_{22.}$		$n_{2j}\ Y_{2j.}\ X_{2j.}$		$n_{2r}\ Y_{2r.}\ X_{2r.}$	$n_{2.}\ Y_{2..}\ X_{2..}$
.							
.							
.							
i	$n_{i1}\ Y_{i1.}\ X_{i1.}$	$n_{i2}\ Y_{i2.}\ X_{i2.}$		$n_{ij}\ Y_{ij.}\ X_{ij.}$		$n_{ir}\ Y_{ir.}\ X_{ir.}$	$n_{i.}\ Y_{i..}\ X_{i..}$
.							
.							
.							
v	$n_{v1}\ Y_{v1.}\ X_{v1.}$	$n_{v2}\ Y_{v2.}\ X_{v2.}$		$n_{vj}\ Y_{vj.}\ X_{vj.}$		$n_{vr}\ Y_{vr.}\ X_{vr.}$	$n_{v.}\ Y_{v..}\ X_{v..}$
Total	$n_{.1}\ Y_{.1.}\ X_{.1.}$	$n_{.2}\ Y_{.2.}\ X_{.2.}$		$n_{.j}\ Y_{.j.}\ X_{.j.}$		$n_{.r}\ Y_{.r.}\ X_{.r.}$	$n_{..}\ Y_{...}\ X_{...}$

In the above equations

$$n_{i.} = \sum_{j=1}^{r} n_{ij} , \qquad n_{.j} = \sum_{i=1}^{v} n_{ij} , \qquad n_{..} = \sum_{i=1}^{v} \sum_{j=1}^{r} n_{ij}$$

and n_{ij} $(= 0, 1, 2, \cdots)$ = number of individuals in the ijth subclass.

The $1 + v + r + b$ normal equations may be reduced to $v + b$ equations involving only the unknowns $\hat{\tau}_i$ and $\hat{\beta}_g$. Further reduction of the $v + b$ equations to v equations involving only the $\hat{\tau}_i$ and the observations is possible, but not too desirable computationally. The $\hat{\beta}_g$ values may be obtained from an analysis of covariance table similar to Table 2 [see 16, sec. 13.7], and then a set of v equations in $\hat{\tau}_i$ may be obtained, the kth equation being:

$$n_{k.}\hat{\tau}_k - \sum_{j=1}^{r} \frac{n_{kj}}{n_{.j}} \sum_{i=1}^{v} n_{ij}\hat{\tau}_i = Y_{k..} - \sum_{j=1}^{r} n_{kj}\bar{y}_{.j.} - \sum_{g=1}^{b} \hat{\beta}_g \left\{ X_{gk..} - \sum_{j=1}^{r} n_{kj}\bar{x}_{g.j.} \right\} = Q_{i.} . \tag{I-7}$$

The v equations from (I 7) and the equation $\sum_{i=1}^{r} \hat{\tau}_i = 0$ yield unique values for the $\hat{\tau}_i$. With solutions for the $\hat{\tau}_i$, and with the additional equation $\sum_{j=1}^{r} \hat{\beta}_j = 0$, $\hat{\mu}$ and $\hat{\beta}_j$ may now be evaluated; the adjusted means are $\hat{\mu} + \hat{\tau}_i$ and $\hat{\mu} + \hat{\beta}_j$ for the first and second classifications respectively.

In the event that formula (I–1) is appropriate, an estimate of β is:

$$\hat{\beta} = \frac{\sum_i \sum_j \sum_h Y_{ijh}X_{ijh} - \sum_j Y_{.j.}\bar{x}_{.j.}}{\sum_i \sum_j \sum_h X^2_{ijh} - \sum X^2_{.j.}/n_{.j} = W_{xx}} - \frac{\sum_i \hat{\tau}_i(X_{i..} - \sum_j n_{ij}\bar{x}_{.j.})}{W_{xx}} \tag{I-8}$$

Likewise, the kth equation of the v equations in $\hat{\tau}_i$ is:

$$n_{k.}\hat{\tau}_k - \sum_i \frac{n_{kj}}{n_{.j}} \sum_i n_{ij}\hat{\tau}_i - \frac{X_{k..} - \sum_j n_{kj}\bar{x}_{.j.}}{W_{xx}} \sum_i \hat{\tau}_i(X_{i..} - \sum_j n_{ij}\bar{x}_{.j.})$$
$$= Y_{k..} - \sum_j n_{kj}\bar{y}_{.j.} - \frac{(X_{k..} - \sum_j n_{kj}\bar{x}_{.j.})}{W_{xx}} \cdot (\sum_i \sum_j \sum_h Y_{ijh}X_{ijh} - \sum_j Y_{.j.}\bar{x}_{.j.}). \tag{I-9}$$

The above v equations plus the equation $\sum_{i=1}^{v} \hat{\tau}_i = 0$ result in unique

solutions for the $\hat{\tau}_i$. The estimate of $\hat{\beta}$ is then obtained from formula (I–8) or $\hat{\beta}$ may be obtained directly from Table 2 as $\hat{\beta} = D_{xy}/D_{xx}$.

If only a variance analysis is to be performed, then each $X_{gijh} - \bar{x}_g$ is set equal to zero in equations (I-1) to (I-7). In obtaining the sums of squares given in Table 2, analyses of variance are obtained for the X_g and the Y variates and their cross products. Estimates of μ, τ_i and ρ_j , obtained from equations (I-1) to (I-7) with each $(X_{gijh} - \bar{x}_g)$ set equal to zero, will be denoted as μ' , ρ_j' , and τ_i' . Estimates of μ and ρ_j obtained from equations (I-2), (I-3), (I-5), and $\sum_{j=1}^{r} \hat{\rho}_j = 0$ when each $\hat{\tau}_i$ and each $(X_{gijh} - \bar{x}_g)$ are set equal to zero will be denoted as μ'', and ρ_j'' . Likewise, the estimate of μ and τ_i obtained from equations (I-2), (I-3), (I-4), and $\sum_{i=1}^{v} \hat{\tau}_i = 0$ when each $\hat{\rho}_j$ and each $(X_{gijh} - \bar{x}_g)$ are set equal to zero will be designated as μ^* and τ_i^* . The estimate of μ obtained from (I-2) and (I-3) when each $\hat{\tau}_i$, each $\hat{\rho}_j$, and each $(X_{gijh} - \bar{x}_g)$ are set equal to zero is $\mu''' = \bar{y}$.

TABLE 2

ANALYSIS OF COVARIANCE (LINEAR) FOR A TWO-WAY CLASSIFICATION WITH UNEQUAL NUMBERS IN THE SUBCLASSES—CASE I.

Source of variation	D.f.	Cross products y^2	xy	x^2
Total (corrected for mean)	$n_{..} - 1$	T_{yy}	T_{xy}	T_{xx}
Second factor (ignoring 1st factor)	$r - 1$	R_{yy}	R_{xy}	R_{xx}
First factor (eliminating 2nd factor)	$v - 1$	V_{yy}	V_{xy}	V_{xx}
Residual	$n_{..} - r - v + 1$	D_{yy}	D_{xy}	D_{xx}
Second factor (eliminating 1st factor)	$r - 1$	B_{yy}	B_{xy}	B_{xx}

	Adjusted sums of squares*	
	D.f.	*Ss*
Residual	$n_{..} - r - v$	$D'_{yy} = D_{yy} - D_{xy}^2/D_{xx}$
First factor + residual	$n_{..} - r - 1$	$W'_{yy} = W_{yy} - W_{xy}^2/W_{xx}$
First factor (eliminating regression and second factor)	$v - 1$	$V'_{yy} = W'_{yy} - D'_{yy}$
Second factor + residual	$n_{..} - v - 1$	$U'_{yy} = U_{yy} - U_{xy}^2/U_{xx}$
Second factor (eliminating regression and first factor)	$r - 1$	$B'_{yy} = U'_{yy} - D'_{yy}$

*$W_{yy} = V_{yy} + D_{yy}$; $W_{xy} = V_{xy} + D_{xy}$; $W_{xx} = V_{xx} + D_{xx}$.
$U_{yy} = B_{yy} + D_{yy}$; $U_{xy} = B_{xy} + D_{xy}$; $U_{xx} = B_{xx} + D_{xx}$.

In computing the sums of squares for the variance analysis, the following procedure is used:*

Total sum of squares $-SS(\mu''')$

$$= \sum_{i=1}^{v} \sum_{j=1}^{r} \sum_{h=1}^{n_{ij}} Y_{ijh}^2 - Y_{...}^2/n_{..} = T_{yy} . \tag{I-10}$$

$$SS(\mu'', \rho_i'') - SS(\mu''')$$

$$= (\mu'' - \mu''')Y_{...} + \sum_j \rho_j'' Y_{.j.} = \sum_j \frac{Y_{.j.}^2}{n_{.j}} - \frac{Y_{...}^2}{n_{..}} = R_{yy} . \tag{I-11}$$

$$SS(\mu^*, \tau_i^*) - SS(\mu''')$$

$$= (\mu^* - \mu''')Y_{...} + \sum_i \tau_i^* Y_{i..} = \sum_i \frac{Y_{i..}^2}{n_{i.}} - \frac{Y_{...}^2}{n_{..}} = A_{yy} . \tag{I-12}$$

$$SS(\mu', \tau_i', \rho_i') - SS(\mu'', \rho_i'')$$

$$= (\mu' - \mu'')Y_{...} + \sum_j (\rho_j' - \rho_j'')Y_{.j.} + \sum_i \tau_i' Y_{i..}$$

$$= \mu' Y_{...} + \sum_j \rho_j' Y_{.j.} + \sum_i \tau_i' Y_{i..} - \sum_j \frac{Y_{.j.}^2}{n_{.j}} \tag{I-13}$$

$$= \sum_{i=1}^{v} \tau_i' \left(Y_{i..} - \sum_{j=1}^{r} n_{ij} \bar{y}_{.j.} \right) = V_{yy} .$$

$$SS(\mu', \rho_j', \tau_i') - SS(\mu^*, \tau_i^*)$$

$$= \mu' Y_{...} + \sum_j \rho_j' Y_{.j.} + \sum_i \tau_i' Y_{i..} - \sum_i Y_{i..}^2/n_{i.} \tag{I-14}$$

$$= \sum_{j=1}^{r} \rho_j' \left(Y_{.j.} - \sum_{i=1}^{v} n_{ij} \bar{y}_{i..} \right) = B_{yy} .$$

Total $- SS(\mu', \rho_j', \tau_i')$

$$= \sum_{i=1}^{v} \sum_{j=1}^{r} \sum_{h=1}^{n_{ij}} Y_{ijh}^2 - \mu' Y_{...} - \sum_{j=1}^{r} \rho_j' Y_{.j.} - \sum_i \tau_i' Y_{i..} = D_{yy} . \tag{I-15}$$

The sums of squares for a classification eliminating the effects of regression, the mean, and the other classification are given in Table 2. Alternatively, the sum of squares for the first classification adjusted for the other effects is $V_{yy}' = SS(\hat{\mu}, \hat{\tau}_i, \hat{\rho}_i, \hat{\beta}) - SS(\mu^+, \rho_j^+,$ and $\beta^+)$, where μ^+, ρ_j^+, and β^+ are obtained from the normal equations with each $\hat{\tau}_i$ set equal to zero.

The various sums of squares for the X variate are obtained in a

*$SS(\mu'', \rho''j)$ is the sum of squares attributable to μ'' and the ρ''_j; etc.

similar manner with μ_x' , τ_{x_i}' , ρ_{x_j}' , μ_x'' , etc. being the corresponding estimates of the effects for the X variate.

The various sums of cross products for the linear model given by formula (I–1) are computed as follows:

$$T_{xy} = \sum_{i=1}^{p} \sum_{j=1}^{r} \sum_{h=1}^{n_{ij}} X_{ijh} Y_{ijh} - \frac{X_{\dots} Y_{\dots}}{n_{\cdot\cdot}} ; \qquad \text{(I–16)}$$

$$R_{x_j} = \sum_{j=1}^{r} \frac{X_{\cdot j \cdot} Y_{\cdot j \cdot}}{n_{\cdot j \cdot}} - \frac{X_{\dots} Y_{\dots}}{n_{\cdot\cdot}} ; \qquad \text{(I–17)}$$

$$A_{xy} = \sum_{i=1}^{p} \frac{X_{i\dots} Y_{i\dots}}{n_{i\cdot}} - \frac{X_{\dots} Y_{\dots}}{n_{\cdot\cdot}} ; \qquad \text{(I–18)}$$

$$\begin{aligned} V_{xy} &= \sum_{i=1}^{p} \tau_{x_i}' \left\{ Y_{i\dots} - \sum_{j=1}^{r} n_{ij} \bar{y}_{\cdot j} \right\} \\ &= \sum_{i=1}^{p} \tau_i' \left\{ X_{i\dots} - \sum_{j} n_{ij} \bar{x}_{\cdot j \cdot} \right\} \\ &= \mu' X_{\dots} + \sum \rho_j' X_{\cdot j \cdot} + \sum \tau_i' X_{i\dots} - \sum X_{\cdot j \cdot} \bar{y}_{\cdot j} \\ &= \mu_x' Y_{\dots} + \sum \rho_{x_j}' Y_{\cdot j \cdot} + \sum \tau_{x_i}' Y_{i\dots} - \sum X_{\cdot j \cdot} \bar{y}_{\cdot j} ; \end{aligned} \qquad \text{(I–19)}$$

$$D_{xy} = T_{xy} - R_{xy} - V_{xy} = T_{xy} - B_{xy} - A_{xy} ; \qquad \text{(I–20)}$$

The above formulae for sums of products are applicable for the gth X variate in a multiple covariance set-up.

If the experiment is designed as a three-way classification with covariates the linear model is:*

$$Y_{ijhf} = \mu + \alpha_i + \gamma_j + \delta_h + \epsilon_{ijhf} + \sum_{g=1}^{b} \beta_g (X_{gijhf} - \bar{x}_g), \qquad \text{(I–21)}$$

where μ = a common mean effect, α_i = effect common to the ith level of the first or A classification, γ_j = effect common to the jth level of the second or C classification, δ_h = effect common to the hth level of the third or D classification, ϵ_{ijhf} = a random error effect, and β_g = true partial regression of Y on X_g estimated from the residual line in the analysis of covariance. There are no interaction terms for a Case I analysis. The normal equations for the various effects ($\hat{\mu}$, $\hat{\alpha}_i$, γ_j , $\hat{\delta}_h$, and $\hat{\beta}_g$, respectively) are:

*This model holds for either a complete or incomplete factorial arrangement of the combinations. For an incomplete factorial arrangement (*e.g.* the latin square) care must be exercised in summing over the various subscripts.

$$n_{\ldots}\hat{\mu} + \sum_{i=1}^{a} n_{i\ldots}\hat{\alpha}_i + \sum_{j=1}^{c} n_{.j.}\hat{\gamma}_j + \sum_{h=1}^{d} n_{\ldots h}\hat{\delta}_h = Y_{\ldots\ldots} ; \qquad \text{(I–22)}$$

$$n_{i\ldots}(\hat{\mu} + \hat{\alpha}_i) + \sum_{j} n_{ij.}\hat{\gamma}_j + \sum_{h} n_{i\ldots h}\hat{\delta}_h + \sum_{g} \hat{\beta}_g(X_{gi\ldots} - n_{i\ldots}\bar{x}_g) = Y_{i\ldots\ldots} ; \qquad \text{(I–23)}$$

$$n_{.j.}(\hat{\mu} + \hat{\gamma}_j) + \sum_{i} n_{ij.}\hat{\alpha}_i + \sum_{h} n_{.j\ldots h}\hat{\delta}_h + \sum_{g} \hat{\beta}_g(X_{g.j\ldots} - n_{.j.}\bar{x}_g) = Y_{.j\ldots\ldots} ; \qquad \text{(I–24)}$$

$$n_{\ldots h}(\hat{\mu} + \hat{\delta}_h) + \sum_{i} n_{i\ldots h}\hat{\alpha}_i + \sum_{j} n_{.j\ldots h}\hat{\gamma}_j + \sum_{g} \hat{\beta}_g(X_{g\ldots h.} - n_{\ldots h}\bar{x}_g) = Y_{\ldots\ldots h.} ; \qquad \text{(I–25)}$$

$$\sum_{i} \hat{\alpha}_i(X_{gi\ldots} - n_{i\ldots}\bar{x}_g) + \sum_{j} \hat{\gamma}_j(X_{g.j\ldots} - n_{.j.}\bar{x}_g) + \sum_{h} \hat{\delta}_h(X_{g\ldots h.} - n_{\ldots h}\bar{x}_g) + \sum_{i}\sum_{j}\sum_{h}\sum_{f} (X_{gijhf} - \bar{x}_g) \sum_{g} \hat{\beta}_g(X_{gijhf} - \bar{x}_g) = \sum_{i}\sum_{j}\sum_{h}\sum_{f=1}^{n_{ijh}} Y_{ijhf}(X_{gijhf} - \bar{x}_g). \qquad \text{(I–26)}$$

The above equations plus the equations

$$\sum_{i=1}^{a} \hat{\alpha}_i = \sum_{j=1}^{c} \hat{\gamma}_j = \sum_{h=1}^{d} \hat{\delta}_h = 0$$

yield unique estimates of the effects.

The form for the linear covariance analysis is indicated in Table 3. In general, the various sums of squares for the variance analysis are obtained as indicated in Table 3. The procedure is easily generalized for a q-way classification with b covariates. The algebra and the arithmetic become more difficult but the principles are the same.

In order to obtain a test of significance for two treatment means, it is necessary to compute the variances for the two means considered. Although the variances and covariances of the $Q_{i.}$ values in (I–7) are known [12], it is rather cumbersome to compute the variance of a difference for two effects, say $\hat{\tau}_1$ and $\hat{\tau}_2$, but if this is desired, Rao [15] has described the general procedure. Also, one could obtain an average coefficient for the variance of a treatment mean and then use the resulting average standard error of a mean in one of the multiple range tests [see 7, ch. II].

TABLE 3

ANALYSIS OF COVARIANCE (LINEAR) FOR A THREE-WAY CLASSIFICATION WITH UNEQUAL NUMBERS IN THE SUBCLASSES—CASE I*

Source of variation	D.f.	Sums of products y^2	xy	x^2
Total (eliminating mean)	$n_{\cdots} - 1$	T_{yy}	T_{xy}	T_{xx}
A (ignoring C and D; elim. mean)	$a - 1$	U_{yy}	U_{xy}	U_{xx}
C (ign. D; elim. A and mean)	$c - 1$	V_{yy}	V_{xy}	V_{xx}
D (elim. A, C, and mean)	$d - 1$	D_{yy}	D_{xy}	D_{xx}
Residual (by subtraction)	f_r	R_{yy}	R_{xy}	R_{xx}
C(elim. A, D, and mean)	$c - 1$	C_{yy}	C_{xy}	C_{xx}
A (elim. C, D, and mean)	$a - 1$	A_{yy}	A_{xy}	A_{xx}
Residual (elim. regression, A, C, D, and mean)	$f_r - 1$	$R'_{yy} = R_{yy} - R^2_{xy}/R_{xx}$		
Residual + D	$f_r + d - 2$	$S'_{yy} = R_{yy} + D_{yy} - (R_{xy} + D_{xy})^2/(R_{xx} + D_{xx})$		
D (elim. regression, A, C, and mean)	$d - 1$	$D'_{yy} = S'_{yy} - R'_{yy}$		
C + residual	$f_r + c - 2$	$W'_{yy} = R_{yy} + C_{yy} - (R_{xy} + C_{xy})^2/(R_{xx} + C_{xx})$		
C (elim. regression, A, D, and mean)	$c - 1$	$C'_{yy} = W'_{yy} - R'_{yy}$		
A + residual	$f_r + a - 2$	$Z'_{yy} = A_{yy} + R_{yy} - (A_{xy} + R_{xy})^2/(A_{xx} + R_{xx})$		
A (elim. regression, C, D, and mean)	$a - 1$	$A'_{yy} = Z'_{yy} - R'_{yy}$		

*$T_{yy} = \sum\sum\sum\sum Y^2_{ijhf} - SS(\mu''')$; $U_{yy} = SS(\mu^* \alpha_i^*) - SS(\mu''')$;
$V_{yy} = SS(\mu^-, \alpha_{\bar{i}}, \gamma_{\bar{j}}) - SS(\mu^*, \alpha_i^*)$; $D_{yy} = SS(\mu', \alpha'_i, \gamma'_j, \delta'_h) - SS(\mu^-, \alpha_{\bar{i}}, \gamma_{\bar{j}})$;
$D'_{yy} = SS(\hat{\mu}, \hat{\alpha}_i, \hat{\gamma}_j, \hat{\delta}_h, \hat{\beta}_g) - SS(\mu'', \alpha_i'', \gamma_j'', \beta_g'')$; $R_{yy} = T_{yy} - U_{yy} - V_{yy} - D_{yy}$

CASE II

When interaction is present in a two-way classification with b covariates, the linear model is:

$$Y_{ijh} = \mu + \tau_i + \rho_j + \rho\tau_{ij} + \epsilon_{ijh} + \sum_{g=1}^{b} \beta_g(X_{gijh} - \bar{x}_g), \qquad \text{(II-1)}$$

where μ, τ_i , ρ_j , and β_g are defined in equation (I-2) and where $\rho\tau_{ij}$ = an interaction effect common to the ijth combination of the two categories. The effects are considered to be fixed effects and $n_{ij} = 1, 2, \cdots$ (i.e., $n_{ij} > 0$).

The normal equations for $\hat{\mu}$, $\hat{\tau}_i$, $\hat{\rho}_j$, $\widehat{\rho\tau}_{ij}$, and $\hat{\beta}_g$, respectively, are:

$$n_{..}\hat{\mu} + \sum_{j=1}^{r} n_{.j}\hat{\rho}_j + \sum_{i=1}^{v} n_{i.}\hat{\tau}_i + \sum_{i=1}^{v}\sum_{j=1}^{r} n_{ij}\widehat{\rho\tau}_{ij} = Y_{...}\ ; \qquad \text{(II-2)}$$

$$n_{i.}(\hat{\mu} + \hat{\tau}_i) + \sum_{j=1}^{r} n_{ij}(\hat{\rho}_j + \widehat{\rho\tau}_{ij}) + \sum_{g=1}^{b} \hat{\beta}_g(X_{gi..} - n_{i.}\bar{x}_g) = Y_{i..}\ ; \qquad \text{(II-3)}$$

$$n_{.j}(\hat{\mu} + \hat{\rho}_j) + \sum_{i=1}^{v} n_{ij}(\hat{\tau}_i + \widehat{\rho\tau}_{ij}) + \sum_{g=1}^{b} \hat{\beta}_g(X_{g.j.} - n_{.j}\bar{x}_g) = Y_{.j.}\ ; \qquad \text{(II-4)}$$

$$n_{ij}(\hat{\mu} + \hat{\rho}_j + \hat{\tau}_i + \widehat{\rho\tau}_{ij}) + \sum_{g=1}^{b} \hat{\beta}_g(X_{gij.} - n_{ij}\bar{x}_g) = Y_{ij.}\ ; \qquad \text{(II-5)}$$

$$\begin{aligned}\sum_i \hat{\tau}_i(X_{gi..} - n_{i.}\bar{x}_g) + \sum_j \hat{\rho}_j(X_{g.j.} - n_{.j}\bar{x}_g) \\ + \sum_i \sum_j \widehat{\rho\tau}_{ij}(X_{gij.} - n_{ij}\bar{x}_g) \\ + \sum_i \sum_j \sum_h (X_{gijh} - \bar{x}_g) \sum_g \hat{\beta}_g(X_{gijh} - \bar{x}_g) \\ = \sum_i \sum_j \sum_h Y_{ijh}(X_{gijh} - \bar{x}_g).\end{aligned} \qquad \text{(II-6)}$$

The above $1 + v + r + rv + b$ equations plus the following result in unique solutions for the $\hat{\mu}$, $\hat{\rho}_j$, $\hat{\tau}_i$, $\widehat{\rho\tau}_{ij}$, and $\hat{\beta}_g$:

$$\sum_{i=1}^{v} \hat{\tau}_i = \sum_{j=1}^{r} \hat{\rho}_j = \sum_{i=1}^{v} \widehat{\rho\tau}_{ij} = \sum_{j=1}^{r} \widehat{\rho\tau}_{ij} = 0. \qquad \text{(II-7)}$$

If $b = 1$, then an estimate of $\beta = \beta_1$ is obtained from the within subclasses line in the analysis of covariance (Table 4) as:

$$\hat{\beta} = \frac{\sum_{i=1}^{v}\sum_{j=1}^{r}\left\{\sum_{h=1}^{n_{ij}} X_{ijh}Y_{ijh} - X_{ij.}Y_{ij.}/n_{ij}\right\} = S_{xy}}{\sum_{i=1}^{v}\sum_{j=1}^{r}\left\{\sum_{h=1}^{n_{ij}} X_{ijh}^2 - X_{ij.}^2/n_{ij}\right\} = S_{xx}}. \qquad \text{(II-8)}$$

If $b > 1$ then the various $\hat{\beta}_g$ may be obtained from the following b equations [see 16, sec. 13.7]:

$$\begin{aligned}\hat{\beta}_1 \sum_i \sum_j \sum_h (X_{gijh} - \bar{x}_{gij.})(X_{1ijh} - \bar{x}_{1ij.}) \\ + \hat{\beta}_2 \sum_i \sum_j \sum_h (X_{gijh} - \bar{x}_{gij.})(X_{2ijh} - \bar{x}_{2ij.}) \\ + \cdots + \hat{\beta}_b \sum_i \sum_j \sum_h (X_{gijh} - \bar{x}_{gij.})(X_{bijh} - \bar{x}_{bij.}) \\ = \sum_i \sum_j \sum_h (Y_{ijh} - \bar{y}_{ij.})(X_{gijh} - \bar{x}_{gij.}),\end{aligned} \tag{II-9}$$

for $g = 1, 2, \cdots, b$.

With the above estimates for $\hat{\beta}_g$, the remaining effects are estimated as follows:

$$\hat{\mu} = \frac{1}{rv} \sum_{i=1}^{v} \sum_{j=1}^{r} (\bar{y}_{ij.} - \sum_g \hat{\beta}_g(\bar{x}_{gij.} - \bar{x}_g)); \tag{II-10}$$

$$\hat{\mu} + \hat{\tau}_i = \frac{1}{r} \sum_{j=1}^{r} (\bar{y}_{ij.} - \sum_g \hat{\beta}_g(\bar{x}_{gij.} - \bar{x}_g)); \tag{II-11}$$

$$\hat{\mu} + \hat{\rho}_j = \frac{1}{v} \sum_{i=1}^{v} (\bar{y}_{ij} - \sum_g \hat{\beta}_g(\bar{x}_{gij.} - \bar{x}_g)); \tag{II-12}$$

$$\widehat{\rho\tau}_{ij} = \bar{y}_{ij.} - \hat{\mu} - \hat{\rho}_j - \hat{\tau}_i - \sum_g \hat{\beta}_g(\bar{x}_{gij.} - \bar{x}_g). \tag{II-13}$$

If one or more of the $n_{ij} = 0$, no analysis is possible unless a zero (or some other constant value) is inserted for the corresponding $\widehat{\rho\tau}_{ij}$; this results in the following changes for formula (II-7) [8, 10]:

$$\sum_{i=1}^{vj} \widehat{\rho\tau}_{ij} = \sum_{j=1}^{ri} \widehat{\rho\tau}_{ij} = 0; \tag{II-14}$$

These changes result in a biased analysis. The size of the bias depends upon the actual values of $\rho\tau_{ij}$ which are estimated as zero and upon the number of n_{ij} equal to zero in the analysis. When some $n_{ij} = 0$, possible recourses are to use the biased Case II analysis, to use a Case I analysis assuming no interaction, or to use an among subclasses and within subclasses analysis.

Before computing the various sums of products it should be noted that there may be no interest in discussing the category means when interaction is assumed present. If there is interest in the category means then weighted means may be more appropriate than the unweighted means. If an among and a within groups variance or co-

variance analysis is suitable (i.e., the rv or $r_{\cdot} = \sum_i r_i = \sum_j v_j = v_{\cdot}$ subclass means are to be compared), then there is little difficulty in computing the among and within subclasses analysis and in comparing the subclass means, $\bar{y}_{ij\cdot}$, either by an F test or a multiple range test [see 7, ch. 11]*.

If tests of significance are required for category means or if components of variance are to be estimated by Henderson's Method 3 [11], the various sums of products in Table 4 need to be computed. The first three rows of sums of products in Table 4 are obtained from a Case I analysis (i.e., formulae (I–10), (I–11), (I–13), (I–16), (I–17), and (I 19). The within subclasses sums of products are obtained in the usual manner [see formula (II 8)]. The interaction sums of products are computed as:

$$I_{yy} = SS(\bar{\mu}, \bar{\rho}_j, \bar{\tau}_i, \overline{\rho\tau}_{ij}) - SS(\mu', \rho_j', \tau_i')$$

$$= \sum_i \sum_j \frac{Y_{ij\cdot}^2}{n_{ij}} - \mu' Y_{\cdots} - \sum \rho_j' Y_{\cdot j\cdot} - \sum \tau_i' Y_{i\cdot\cdot} ; \qquad \text{(II–15)}$$

$$I_{xy} = \sum_i \sum_j \frac{X_{gij\cdot} Y_{ij\cdot}}{n_{ij\cdot}} - \mu' X_{g\cdots} - \sum \rho_j' X_{g\cdot j\cdot} - \sum \tau_i' X_{gi\cdot\cdot} \qquad \text{(II–16)}$$

$$= \sum_i \sum_j \frac{X_{gij\cdot} Y_{ij\cdot}}{n_{ij}} - \mu_x' Y_{\cdots} - \sum \rho_{xj}' Y_{\cdot j\cdot} - \sum \tau_{xi}' Y_{i\cdot\cdot} ;$$

$$I_{xx} = \sum_i \sum_j \frac{X_{gij\cdot}^2}{n_{ij}} - \mu_x' X_{g\cdots} - \sum \rho_{xj}' X_{g\cdot j\cdot} - \sum \tau_{xi}' X_{gi\cdot\cdot} , \qquad \text{(II–17)}$$

where the estimates μ', ρ_j', and τ_i' are defined in the paragraph preceding formula (I–10) and where $\bar{\mu}$, $\bar{\rho}_j$, $\bar{\tau}_i$, and $\widehat{\rho\tau}_{ij}$ are the estimates obtained from formulae (II-1) to (II–7) when each $X_{gijh} - \bar{x}_g$ is set equal to zero.

The sums of products for the first factor (eliminating mean, second factor and interaction) are [10, 21]:

---

*In a multiple range test, the standard error of a mean used to compute a significant range for comparing two means, say $\bar{y}_{1\cdot\cdot}$ and $\bar{y}_{2\cdot\cdot}$ with $n_{1\cdot}$ and $n_{2\cdot}$ observations, respectively, is (D. B. Duncan, written correspondence):

$$\sqrt{\frac{S_{yy}}{n_{\cdot\cdot} - r_{\cdot}} \left\{ \frac{1}{2} \left(\frac{1}{n_{1\cdot}} + \frac{1}{n_{2\cdot}} \right) \right\}}$$

In a covariance analysis, the average variance of a difference for A category means, say i and i', is approximately equal to

$$\sqrt{\frac{S'_{yy}}{n_{\cdot\cdot} - r_{\cdot} - 1} \left\{ \frac{1}{2} \left(\frac{1}{n_{i\cdot}} + \frac{1}{n_{i'\cdot}} \right) \left(1 + \frac{A_{xx}}{S_{xx}(v - 1)} \right) \right\}}.$$

$$A_{yy} = SS(\bar{\mu}, \bar{\rho}_i, \bar{\tau}_i, \bar{\rho}\bar{\tau}_{ij}) - SS(\mu^+, \rho_i^+, \rho\tau_{ij}^+)$$
$$= \sum w_{i\cdot} \bar{y}_{i\cdot}^2 - \frac{(\sum w_{i\cdot}\bar{y}_{i\cdot})^2}{\sum w_{i\cdot}}, \qquad \text{(II–18)}$$

where μ^+, ρ_i^+, $\rho\tau_{ij}^+$ are the estimates obtained from the normal equations when each $\hat{\tau}_i$ and each $(X_{gijh} - \bar{x}_g)$ is set equal to zero;

$$A_{xy} = \sum_i w_{i\cdot}\bar{x}_{i\cdot}\bar{y}_{i\cdot} - \frac{\sum w_{i\cdot}\bar{x}_{i\cdot} \sum w_{i\cdot}\bar{y}_{i\cdot}}{\sum w_{i\cdot}}; \qquad \text{(II–19)}$$

$$A_{xx} = \sum_i w_{i\cdot}\bar{x}_{i\cdot}^2 - \frac{(\sum w_{i\cdot}\bar{x}_{i\cdot})^2}{\sum w_{i\cdot}}. \qquad \text{(II–20)}$$

TABLE 4

ANALYSIS OF COVARIANCE (LINEAR) FOR A TWO-WAY CLASSIFICATION WITH UNEQUAL NUMBERS IN THE SUBCLASSES—CASE II
(FIRST FACTOR = A; SECOND FACTOR = B).

Source of variation	D.f.	Sums of products y^2	xy	x^2
Total (eliminating mean)	$n_{\cdot\cdot} - 1$	T_{yy}	T_{xy}	T_{xx}
B (elim. mean; ign. A and $A \times B$)	$r - 1$	R_{yy}	R_{xy}	R_{xx}
A (elim. B and mean; ign. $A \times B$)	$v - 1$	V_{yy}	V_{xy}	V_{xx}
$A \times B$ (elim. A, B, and mean)	$(r-1)(v-1)$	I_{yy}	I_{xy}	I_{xx}
Within subclasses $= S$	$n_{\cdot\cdot} - rv$	S_{yy}	S_{xy}	S_{xx}
B (elim. A, $A \times B$, and mean)	$r - 1$	B_{xx}	B_{xy}	B_{xx}
A (elim. B, $A \times B$, and mean)	$v - 1$	A_{yy}	A_{xy}	A_{xx}
S (adj. for regression)	$n_{\cdot\cdot} - rv - 1$	$S'_{yy} = S_{yy} - S_{xy}^2/S_{xx}$		
$S + B$	$n_{\cdot\cdot} - rv + r - 2$	$U'_{yy} = B_{yy} + S_{yy} - (B_{xy} + S_{xy})^2/(B_{xx} + S_{xx})$		
B (adj. for other effects)	$r - 1$	$B'_{yy} = U'_{yy} - S'_{yy}$		
$S + A$	$n_{\cdot\cdot} - rv + v - 2$	$W'_{yy} = A_{yy} + S_{yy} - (A_{xy} + S_{xy})^2/(A_{xx} + S_{xx})$		
A (adj. for other effects)	$v - 1$	$A'_{yy} = W'_{yy} - S'_{yy}$		
$S + A \times B$	$n_{\cdot\cdot} - r - v$	$Z'_{yy} = I_{yy} + S_{yy} - (I_{xy} + S_{xy})^2/(I_{xx} + S_{xx})$		
$A \times B$ (adj. for other effects)	$(r-1)(v-1)$	$I'_{yy} = Z'_{yy} - S'_{yy}$		

In the above,

$$\bar{y}_{i..} = \frac{1}{r}\sum_{j=1}^{r} \bar{y}_{ij.}\,, \qquad \bar{x}_{i..} = \frac{1}{r}\sum_{j=1}^{r} \bar{x}_{ij.}\,, \quad \text{and} \quad w_{i.} = \left(\frac{1}{r^2}\sum_{j}\frac{1}{n_{ij}}\right)^{-1}.$$

Due to a theorem given by Yates [22], the F or z test of significance is appropriate for comparing the two mean squares $A_{yy}/(v-1)$ and $S_{yy}/(n_{..} - rv)$. Likewise,

$$F = \frac{A'_{yy}(n_{..} - rv - 1)}{(v-1)S'_{yy}}$$

is a valid test of the null hypothesis of zero effects in a covariance analysis.

Similarly,

$$B_{xy} = \sum_{j} w_{.j}\bar{y}_{.j.}\bar{x}_{.j.} - \frac{\sum_{j} w_{.j}\bar{x}_{.j.}\ \sum_{j} w_{.j}\bar{y}_{.j.}}{\sum_{j} w_{.j}}\,; \qquad \text{(II–21)}$$

where

$$\bar{y}_{.j.} = \frac{1}{v}\sum_{i=1}^{v} \bar{y}_{ij.}\,, \qquad \bar{x}_{.j.} = \frac{1}{v}\sum_{i=1}^{v} \bar{x}_{ij.}\,, \quad \text{and} \quad w_{.j} = \left(\frac{1}{v^2}\sum_{i=1}^{v}\frac{1}{n_{ij}}\right)^{-1}.$$

For a three-way classification with b covariates and interaction effects not assumed to be zero as in (I–21), the linear model is:

$$Y_{ijhf} = \mu + \alpha_i + \gamma_j + \delta_h + \alpha\gamma_{ij} + \alpha\delta_{ih} + \gamma\delta_{jh} + \alpha\gamma\delta_{ijh} + \epsilon_{ijhf} + \sum_{g=1}^{b}\beta_g(X_{gijhf} - \bar{x}_g), \qquad \text{(II–22)}$$

where μ, α_i, γ_j, δ_h, ϵ_{ijhf}, and β_g are defined in (I–21) and where the remaining terms represent interaction effects (fixed) between the three classifications. Here, n_{ijh} = number of observations for the ijhth combination of the three factors must be greater than zero. If any $n_{ijh} = 0$, then the alternatives described for the preceding unbalanced two-way classification may be used. The normal equations for the various effects are set up in the usual manner; e.g., the normal equation for $\hat{\alpha}_i$ is:

$$n_{i..}(\hat{\mu} + \hat{\alpha}_i) + \sum_{j} n_{ij.}(\hat{\gamma}_j + \widehat{\alpha\gamma}_{ij}) + \sum_{h} n_{i.h}(o_h + \widehat{\alpha\delta}_{ih}) + \sum_{j}\sum_{h} n_{ijh}(\widehat{\gamma\delta}_{jh} + \widehat{\alpha\gamma\delta}_{ijh}) + \sum_{g}\hat{\beta}_g(X_{gi...} - n_{i..}\bar{x}_g) = Y_{i...}\,; \qquad \text{(II–23)}$$

TABLE 5

ANALYSIS OF VARIANCE FOR AN UNBALANCED THREE-WAY CLASSIFICATION—CASE II
(FIRST FACTOR = A; SECOND FACTOR = C; THIRD FACTOR = D).

Source of variation	Degrees of freedom	Sum of squares*
Total (elim. mean only)	$n_{...} - 1$	$\sum\sum\sum\sum Y^2_{ijhf} - Y^2_{...}/n_{...} = T - CT$
A (elim. mean ign. other effects)	$a - 1$	$\sum Y^2_{i...}/n_{i..} - CT$
C (elim. mean and A; ign. others)	$c - 1$	$SS(\mu^{(1)}, \alpha_i^{(1)}, \gamma_j^{(1)}) - \sum Y^2_{i...}/n_{i..}$
$A \times C$ (elim. A, C, mean; ign. others)	$(a - 1)(c - 1)$	$\sum_i\sum_j Y^2_{ij..}/n_{ij.} - SS(\mu^{(1)}, \alpha_i^{(1)}, \gamma_j^{(1)})$
D (elim. A, C, $A \times C$, mean; ign. others)	$d - 1$	$SS(\mu^{(2)}, \alpha_i^{(2)}, \gamma_j^{(2)}, \alpha\gamma_{ij}^{(2)}, \delta_h^{(2)}) - \sum_i\sum_j Y^2_{ij..}/n_{ij..} = \text{II} - \text{I}$
$A \times D$ (elim. A, C, $A \times C$, D, mean; ign. $C \times D$, $A \times C \times D$)	$(a - 1)(d - 1)$	$SS(\mu^{(3)}, \alpha_i^{(3)}, \gamma_j^{(3)}, \alpha\gamma_{ij}^{(3)}, \delta_h^{(3)}, \alpha\delta_{ih}^{(3)}) - \text{II} = \text{III} - \text{II}$
$C \times D$ (elim. all effects but $A \times C \times D$)	$(c - 1)(d - 1)$	$SS(\mu^{(4)}, \alpha_i^{(4)}, \gamma_j^{(4)}, \alpha\gamma_{ij}^{(4)}, \delta_h^{(4)}, \alpha\delta_{ih}^{(4)}, \gamma\delta_{jh}^{(4)}) - \text{III} = \text{IV} - \text{III}$
$A \times C \times D$ (elim. all other effects)	$(a - 1)(c - 1)(d - 1)$	$\sum_i\sum_j\sum_h Y^2_{ijh.}/n_{ijh.} - \text{IV}$
Within subclasses	$n_{...} - acd$	$\sum_i\sum_j\sum_h(\sum_f Y^2_{ijhf} - Y^2_{ijh.}/n_{ijh.})$
A (elim. all other effects)	$a - 1$	$\sum_i\sum_j\sum_h Y^2_{ijh.}/n_{ijh.} - SS(\mu^{(6)}, \gamma_j^{(6)}, \alpha\gamma_{ij}^{(6)}, \delta_h^{(6)}, \alpha\delta_{ih}^{(6)}, \gamma\delta_{jh}^{(6)}, \alpha\gamma\delta_{ijh}^{(6)})$
C (elim. all other effects)	$c - 1$	similar to A (elim. all other effects)
$A \times C$ (elim. all other effects)	$(a - 1)(c - 1)$	$\sum\sum\sum Y^2_{ijh.}/n_{ijh.} - SS(\mu^{(5)}, \alpha_i^{(5)}, \gamma_j^{(5)}, \delta_h^{(5)}, \alpha\delta_{ih}^{(5)}, \gamma\delta_{jh}^{(5)}, \alpha\gamma\delta_{ijh}^{(5)})$
D (elim. all other effects)	$d - 1$	similar to A (elim. all other effects)
$A \times D$ (elim. all other effects)	$(a - 1)(d - 1)$	similar to $A \times C$ (elim. all other effects)
$C \times D$ (elim. all other effects)	$(c - 1)(d - 1)$	similar to $A \times C$ (elim. all other effects)

*The superscript in parentheses refers to the estimates obtained from the normal equations setting the effect (s) not included in the term $SS(\quad)$ equal to zero and omitting the normal equations for these effects.

The sums of squares for the various effects except the mean are presented in Table 5. The method of computation for sums of squares which add to the total is given in the upper part of the table. The resulting mean squares from the next to the last line in the upper part and all those from the lower part of Table 5 are used in making tests of significance of the null hypotheses for each of the main effects and interactions. The within subclasses mean square is the error mean square for all F tests.

Covariance analyses are carried out in the same manner as for the two-way classification. For example, the sum of cross products for $A \times C$ (eliminating all other effects) is:

$$\sum_i \sum_j \sum_h \frac{Y_{ijh\cdot} X_{ijh\cdot}}{n_{ijh}} - \{\mu^{(5)} X_{\cdots} + \sum \alpha_i^{(5)} X_{i\cdots} + \sum_j \gamma_j^{(5)} X_{\cdot j\cdot} + \sum \delta_h^{(5)} X_{\cdot\cdot h\cdot} + \sum_i \sum_h \alpha\delta_{ih}^{(5)} X_{i\cdot h\cdot} + \sum_j \sum_h \gamma\delta_{jh}^{(5)} X_{\cdot jh\cdot} + \sum_i \sum_j \sum_h \alpha\gamma\delta_{ijh}^{(5)} X_{ijh} \}, \qquad \text{(II 24)}$$

where the estimates with the superscript $^{(5)}$ are obtained from the remaining normal equations with each $\widehat{\alpha\gamma}_{ij}$ and each $(X_{qijhf} - \bar{x}_q)$ set equal to zero.

With b covariates, solutions for $\hat{\beta}_q$ may be obtained from formulae similar to (II 9), i.e., the within subclasses sums of squares and cross products for a three (or higher)-way classification. Likewise, a direct extension of the results in Tables 4 and 5 results in a variance or covariance analysis for a q-way classification.

CASE III

If one or both categories in a two-way classification are considered to be a sample of levels or treatments from a large population of levels or treatments for the given category, the interaction mean square is used to test hypotheses about mean effects and to adjust the means for variation due to a covariate(s). Also, the interaction effects, as such, may not be of much interest except in estimating the component of variance associated with interaction effects. Since a weighted estimate of the effect would be preferable, for statistical reasons, to an unweighted estimate, the problem of estimating effects is complicated because the weights are unknown. If the weights are estimated from the data, further statistical problems arise. In certain experiments, the experimenter may have little or no control over the number of observations in each subclass and it may be realistic to assume that the linear model is of the following form:

$$Y_{ijh} = \mu + \tau_i + \rho_j + \rho\tau_{ij} + \epsilon_{ijh} + \beta_1(\bar{x}_{ij.} - \bar{x}) + \beta(X_{ijh} - \bar{x}_{ij.}), \tag{III-1}$$

where μ, τ_i, ρ_j, and ϵ_{ijh}, $\bar{x}$, and X_{ijh} are defined in (II-1) except that τ_i and or ρ_j are considered to be random effects, $\bar{x}_{ij.} = \sum_{h=1}^{n_{ij}} X_{ijh}/n_{ij}$, $\rho\tau_{ij}$ = an interaction effect of the ijth combination of the two categories and is a random effect, β_1 = regression coefficient from the interaction line in the covariance analysis, and β = a within subclasses regression coefficient; if $\beta_1 = \beta$ then (III-1) reduces to (II-1) except for the random effects set-up for (III-1) as opposed to the fixed effects assumption in (II-1). For such situations, a Case III analysis may be desired and the weights may have to be estimated from data in the present or in a past experiment. The weights themselves will have a sampling distribution but the experiment would usually be analyzed assuming that the estimated weights were the true weights.

Two sums of squares could be minimized to obtain least squares estimates of the effects. These are:

$$\sum_{i=1}^{p}\sum_{j=1}^{r}\sum_{h=1}^{n_{ij}} (Y_{ijh} - \mu - \tau_i - \rho_j - \rho\tau_{ij} - \beta_1(\bar{x}_{ij.} - \bar{x}) - \beta(X_{ijh} - \bar{x}_{ij.}))^2 \tag{III-2}$$

and

$$\sum_{i=1}^{p}\sum_{j=1}^{r} w_{ij}(\bar{y}_{ij.} - \mu - \tau_i - \rho_j - \beta_1(\bar{x}_{ij.} - \bar{x}))^2, \tag{III-3}$$

where $\bar{y}_{ij.}$ and $\bar{x}_{ij.}$ = subclass means for the two variates. If $w_{ij} = n_{ij}$ in (III-3) where n_{ij} = number of observations in the ijth subclass, minimization of (III-3) results in the same estimates of μ, τ_i, ρ_j, and β_1 obtained from minimizing (III-2). Since this is true, since w_{ij} is not always equal to n_{ij}, and since β is estimated from the within subclasses sums of products according to formula (II-8), the sum of squares in (III-3) is minimized instead of the one in (III-2).

The problem of correct weighting is relatively simple if the true weights are known. Thus, weighting inversely to the variance of $\bar{y}_{ij.}$ results in:

$$w_{ij} = \frac{1}{\sigma^2_{\rho\tau} + \sigma^2_\epsilon/n_{ij}} = \frac{n_{ij}}{n_{ij}\sigma^2_{\rho\tau} + \sigma^2_\epsilon}, \tag{III-4}$$

where $\sigma^2_{\rho\tau}$ and σ^2_ϵ are the variance components associated with interaction and within subclasses, respectively. In practice σ^2_ϵ and $\sigma^2_{\rho\tau}$ are usually unknown. First, consider the two limiting situations:

(i) $\sigma^2_{\rho\tau}$ is large relative to σ^2_ϵ/n_{ij};

(ii) $\sigma^2_{\rho\tau}$ is small relative to σ^2_ϵ/n_{ij}.

In situation (i), one may set $w_{ij} = 1$ for all practical purposes; i.e., an analysis of covariance is performed on the subclass means. In the second situation, w_{ij} is set equal to n_{ij}, and the analysis goes through as described below using the weights w_{ij} [see 8].

The true situation is usually in between the above two limiting situations, and in order to perform an analysis on the data it will be necessary to have reasonably good estimates of $\sigma^2_{\rho\tau}$ and σ^2_{ϵ}. In many situations the estimated variance (or covariance) components will need to be estimated from the data themselves. One method for doing this is illustrated by Federer [8], where Henderson's [11] Method 1 was used. This method assumes that both effects are random effects; if one of the effects is fixed then a bias results [11]. Recourse to Henderson's [11] Methods 2 or 3 may be made if one of the effects is considered to be fixed.

The normal equations for $\hat{\mu}$, $\hat{\tau}_i$, $\hat{\rho}_j$, and $\hat{\beta}_1$, respectively, are:

$$w_{..}\hat{\mu} + \sum_{i=1}^{v} w_{i.}\hat{\tau}_i + \sum_{j=1}^{r} w_{.j}\hat{\rho}_j + \hat{\beta}_1 \sum_{i=1}^{v}\sum_{j=1}^{r} w_{ij}(\bar{x}_{ij.} - \bar{x}) = \sum_{i=1}^{v}\sum_{j=1}^{r} w_{ij}\bar{y}_{ij.}\ ; \tag{III–5}$$

$$w_{i.}(\hat{\mu} + \hat{\tau}_i) + \sum_{j=1}^{r} w_{ij}(\hat{\rho}_j + \hat{\beta}_1(\bar{x}_{ij.} - \bar{x})) = \sum_{j=1}^{r} w_{ij}\bar{y}_{ij.}\ ; \tag{III–6}$$

$$w_{.j}(\hat{\mu} + \hat{\rho}_j) + \sum_{i=1}^{v} w_{ij}(\hat{\tau}_i + \hat{\beta}_1(\bar{x}_{ij.} - \bar{x})) = \sum_{i=1}^{v} w_{ij}\bar{y}_{ij.}\ ; \tag{III–7}$$

$$\hat{\mu}\sum_{i=1}^{v}\sum_{j=1}^{r} w_{ij}(\bar{x}_{ij.} - \bar{x}) + \sum_{i=1}^{v}\hat{\tau}_i\sum_{j=1}^{r} w_{ij}(\bar{x}_{ij.} - \bar{x}) + \sum_{j=1}^{r}\hat{\rho}_j\sum_{i=1}^{v} w_{ij}(\bar{x}_{ij.} - \bar{x}) + \hat{\beta}_1\sum_{i=1}^{v}\sum_{j=1}^{r} w_{ij}(\bar{x}_{ij.} - \bar{x})^2 = \sum_{i=1}^{v}\sum_{j=1}^{r} w_{ij}\bar{y}_{ij.}(\bar{x}_{ij.} - \bar{x}). \tag{III–8}$$

In the above

$$w_{..} = \sum_{i=1}^{v}\sum_{j=1}^{r} w_{ij}\ ; \qquad w_{i.} = \sum_{j=1}^{r} w_{ij}\ ; \qquad w_{.j} = \sum_{i=1}^{v} w_{ij}\ ;$$

w_{ij} = weight for ijth subclass mean; w_{ij} could be zero and the analysis could still be performed. That is, zero observations could be obtained for certain subclasses: If this were a random event, the interaction effect for that subclass would be set equal to its expected value, zero.

Substituting for $\hat{\mu}$, $\hat{\rho}_j$, and $\hat{\beta}_1$ results in v equations in the $\hat{\tau}_i$, the

kth equation being:

$$\begin{aligned}\hat{\tau}_k\Bigg\{w_{k\cdot} - \sum_{j=1}^{r} w_{kj}^2/w_{\cdot j} - \frac{1}{G_{xx}}\left(\sum_{j=1}^{r} w_{kj}\bar{x}_{kj\cdot} - \sum_{j=1}^{r}\frac{w_{kj}}{w_{\cdot j}}\sum_{i=1}^{v} w_{ij}\bar{x}_{ij\cdot}\right)^2\Bigg\}\\ - \sum_{\substack{i=1\\ i\neq k}}^{v}\hat{\tau}_i\Bigg\{\sum_{j=1}^{r}\frac{w_{ij}w_{kj}}{w_{\cdot j}}\\ - \frac{1}{G_{xx}}\left(\sum_{j=1}^{r} w_{kj}\bar{x}_{kj\cdot} - \sum_{j=1}^{r}\frac{w_{kj}}{w_{\cdot j}}\sum_{i=1}^{v} w_{ij}\bar{x}_{ij\cdot}\right)\\ \times\left(\sum_{j=1}^{r}\frac{w_{ij}}{w_{\cdot j}}\sum_{i=1}^{v} w_{ij}\bar{x}_{ij\cdot} - \sum_{j=1}^{r} w_{ij}\bar{x}_{ij\cdot}\right)\Bigg\}\\ = \sum_{j=1}^{r} w_{kj}\bar{y}_{kj\cdot} - \sum_{j=1}^{r}\frac{w_{kj}}{w_{\cdot j}}\sum_{i=1}^{v} w_{ij}\bar{y}_{ij\cdot}\\ - \frac{G_{xy}}{G_{xx}}\left(\sum_{j=1}^{r} w_{kj}\bar{x}_{kj\cdot} - \sum_{j=1}^{r}\frac{w_{kj}}{w_{\cdot j}}\sum_{i=1}^{v} w_{ij}\bar{x}_{ij\cdot}\right),\end{aligned} \tag{III–9}$$

where

$$G_{xy} = \sum_{i=1}^{v}\sum_{j=1}^{r} w_{ij}\bar{x}_{ij\cdot}\bar{y}_{ij\cdot} - \sum_{j=1}^{r}\frac{1}{w_{\cdot j}}\left(\sum_{i=1}^{v} w_{ij}\bar{y}_{ij\cdot}\right)\left(\sum_{i=1}^{v} w_{ij}\bar{x}_{ij\cdot}\right) \tag{III–10}$$

and

$$G_{xx} = \sum_{i=1}^{v}\sum_{j=1}^{r} w_{ij}\bar{x}_{ij\cdot}^2 - \sum_{j=1}^{r}\left(\sum_{i=1}^{v} w_{ij}\bar{x}_{ij\cdot}\right)^2 \Big/ w_{\cdot j}\,. \tag{III–11}$$

Likewise, the k-th equation of the v equations in the $\hat{\tau}_i$ in terms of $\hat{\beta}_1$ and the observations is:

$$\begin{aligned}w_{k\cdot}\hat{\tau}_k - \sum_{j=1}^{r}\frac{w_{kj}}{w_{\cdot j}}\sum_{i=1}^{v} w_{ij}\hat{\tau}_i = \sum_{j=1}^{r} w_{kj}\bar{y}_{kj\cdot} - \sum_{j=1}^{r}\frac{w_{kj}}{w_{\cdot j}}\sum_{i=1}^{v} w_{ij}\bar{y}_{ij\cdot}\\ - \hat{\beta}_1\left(\sum_{j=1}^{r} w_{kj}\bar{x}_{kj\cdot} - \sum_{j=1}^{r}\frac{w_{kj}}{w_{\cdot j}}\sum_{i=1}^{v} w_{ij}\bar{x}_{ij\cdot}\right).\end{aligned} \tag{III–12}$$

The v equations in (III–12) plus the equation $\sum_{i=1}^{v}\hat{\tau}_i = 0$ result in a solution for the $\hat{\tau}_i$. If a variance analysis is desired each $\bar{x}_{ij\cdot}$ in (III–12) is set equal to zero and the resulting estimates τ_i' of the τ_i are obtained. $\hat{\beta}_1$ in the covariance analysis is the interaction sum of cross products divided by the interaction sum of squares for the covariate, i.e. $\hat{\beta}_1 = E_{xy}/E_{xx}$ (Table 6).

The sums of products in the first three rows of Table 6 are obtained in the usual manner [see formulae (I–10), (I 16), and (II 8)]. The remaining sums of the cross products are given below:

$$W_{xy} = \sum_i \sum_j w_{ij}\bar{y}_{ij.}\bar{x}_{ij.} - \frac{(\sum_i \sum_j w_{ij}\bar{x}_{ij.})(\sum_i \sum_j w_{ij}\bar{y}_{ij.})}{w_{..}} \; ; \quad \text{(III–13)}$$

$$R_{xy} = \sum_{j=1}^{r} \frac{(\sum_i w_{ij}\bar{y}_{ij.})(\sum_i w_{ij}\bar{x}_{ij.})}{w_{.j}} - \frac{(\sum_i \sum_j w_{ij}\bar{x}_{ij.})(\sum_i \sum_j w_{ij}\bar{y}_{ij.})}{w_{..}} . \quad \text{(III–14)}$$

$$\begin{aligned} A_{xy} &= \mu' \sum_i \sum_j w_{ij}\bar{x}_{ij.} \\ &\quad + \sum_j \rho'_j(\sum_i w_{ij}\bar{x}_{ij.}) + \sum_i \tau'_i(\sum_j w_{ij}\bar{x}_{ij.}) \\ &\quad - \sum_j (\sum_i w_{ij}\bar{y}_{ij.})(\sum_i w_{ij}\bar{x}_{ij.})/w_{.j} \\ &= \mu'_x \sum_i \sum_j w_{ij}\bar{y}_{ij.} \\ &\quad + \sum_j \rho'_{xj}(\sum_i w_{ij}\bar{y}_{ij.}) + \sum_i \tau'_{xi}(\sum_j w_{ij}\bar{y}_{ij.}) \\ &\quad - \sum_j (\sum_i w_{ij}\bar{y}_{ij.})(\sum_i w_{ij}\bar{x}_{ij.})/w_{.j} \; , \end{aligned} \quad \text{(III–15)}$$

where μ', ρ'_j, and τ'_i are the estimates obtained from the normal equations with each $(\bar{x}_{ij.} - \bar{x})$ set equal to zero; the μ'_x , ρ'_{xj} , and τ'_{xi} are similar estimates obtained for the X variate;

$$E_{xy} = W_{xy} - R_{xy} - A_{wy} \; ; \quad \text{(III–16)}$$

$$\begin{aligned} B_{xy} &= \mu' \sum_i \sum_j w_{ij}\bar{x}_{ij.} \\ &\quad + \sum_j \rho'_j(\sum_i w_{ij}\bar{x}_{ij.}) + \sum_i \tau'_i(\sum_j w_{ij}\bar{x}_{ij.}) \\ &\quad - \sum_i (\sum_j w_{ij}\bar{y}_{ij.})(\sum_j w_{ij}\bar{x}_{ij.})/w_{i.} \; . \end{aligned} \quad \text{(III–17)}$$

The procedure for obtaining the sums of squares eliminating regression is indicated in Table 6.

The procedure for b covariates is a straightforward extension of the results for a Case II analysis. The normal equations given by formulae (III–5) to (III–8) would be altered for the terms involving $\hat{\beta}_1$ and the $\bar{x}_{ij.}$'s. Instead of the term $\hat{\beta}_1(\bar{x}_{ij.} - \bar{x})$ substitute $\sum_{g=1}^{b} \hat{\beta}_{1g}(\bar{x}_{gij.} - \bar{x}_g)$ in (III -1) and make the corresponding changes in formulae (III–5) to (III -12). The $\hat{\beta}_{1g}$ would be obtained from the Error (interaction) line in the multiple covariance analysis as described for a Case II analysis.

For a three-way classification and for the random effects situation for all three categories, additional statistical problems arise in testing

TABLE 6

COVARIANCE (LINEAR) ANALYSIS FOR AN UNBALANCED TWO-WAY CLASSIFICATION—CASE III.

Source of variation	D.f.	Sums of products y^2	xy	x^2
Total (elim. mean only)	$n_{..} - 1$	T_{yy}	T_{xy}	T_{xx}
Within subclasses	$n_{..} - r$	S_{yy}	S_{xy}	S_{xx}
Among subclasses	$r - 1$	by subtraction		
Among weighted subclass totals	$r - 1$	W_{yy}	W_{xy}	W_{xx}
B (elim. mean; ign. A)	$r - 1$	R_{yy}	R_{xy}	R_{xx}
A (elim. mean and B) $= A$	$v - 1$	A_{yy}	A_{xy}	A_{xx}
Error (interaction) $= E$	f_e	E_{yy}	E_{xy}	E_{xx}
B (elim. mean and A) $= B$	$r - 1$	B_{yy}	B_{xy}	B_{xx}
E (elim. A, B, mean, and regression)	$f_e - 1$	$E'_{yy} = E_{yy} - E_{xy}^2/E_{xx}$		
$A + E$	$f_e + v - 2$	$U'_{yy} = A_{yy} + E_{yy} - (A_{xy} + E_{xy})^2/(A_{xx} + E_{xx})$		
A (elim. B, mean, and regression)	$v - 1$	$A'_{yy} = U'_{yy} - E'_{yy}$		
$B + E$	$f_e + r - 2$	$V'_{yy} = B_{yy} + E_{yy} - (B_{xy} + E_{xy})^2/(B_{xx} + E_{xx})$		
B (elim. A, mean, and regression)	$r - 1$	$B'_{yy} = V'_{yy} - E'_{yy}$		

hypotheses about main effects. For example, consider that the linear model is:

$$
\begin{aligned}
Y_{ijhf} = \mu &+ \alpha_i + \gamma_j + \delta_h + \alpha\gamma_{ij} + \alpha\delta_{ih} + \gamma\delta_{jh} + \alpha\gamma\delta_{ijh} + \epsilon_{ijhf} \\
&+ \beta_1(\bar{x}_{ij..} - \bar{x}_{i...} - \bar{x}_{.j..} + \bar{x}) \\
&+ \beta_2(\bar{x}_{i.h.} - \bar{x}_{i...} - \bar{x}_{..h.} + \bar{x}) \\
&+ \beta_3(\bar{x}_{.jh.} - \bar{x}_{.j..} - \bar{x}_{..h.} + \bar{x}) \\
&+ \beta_4(\bar{x}_{ijh.} + 2\bar{x}_{i...} + 2\bar{x}_{.j..} + 2\bar{x}_{..h.} \\
&\qquad - \bar{x}_{ij..} - \bar{x}_{i.h.} - \bar{x}_{.jh.} - 4\bar{x}) \\
&+ \beta(X_{ijhf} - \bar{x}_{ijh.}),
\end{aligned}
\qquad \text{(III-18)}
$$

where the effects are defined in (II–22) except that all effects are random effects, β_1 , β_2 , β_3 and β_4 are different regression coefficients from the various interaction lines in the analysis of covariance. In such a model, there is no single line in the analysis of covariance that is suitable for testing hypotheses about main effects and for adjusting for variation in the covariate. A simplifying assumption would be to assume $\beta_1 = \beta_2 = \beta_3 = \beta_4$ and possibly that $\beta_4 = \beta$ also. This would simplify the covariance problem but would not simplify the hypotheses testing problem [7, ch. VIII]. A simplification in hypothesis testing would be possible if one or more of the two-factor interactions could be assumed relatively small, and then there would be a single interaction mean square which could be used to test hypotheses about main effects concerned.

The normal equations and the Case III covariance analysis for the unbalanced three-way classification represent a straightforward extension of the Case II analysis for a three-way classification and of a Case III analysis for a two-way classification. Solutions for the various effects are obtained as before.

Estimation of variance components from a given set of data, either the present or a past experiment, presents some problems even for an unbalanced two-way classification. As a first approximation in an experiment one could use Henderson's [11] Method 1 assuming both effects random and assuming w_{ij} fixed [see 8]. The process is repeated on each analysis and the new weights are used in the subsequent analysis. The process is repeated until the weights stabilize.

Alternatively, one could use Henderson's [11] Method 3 and obtain the expectation of I'_{vv} in Table 4 as a first estimate for $\sigma^2_{\rho\tau}$. σ^2_ϵ is estimated from the within subclasses mean squares as $\hat{\sigma}^2_\epsilon = S'_{yy}/(n_{..} - rv - 1)$. The analysis of covariance in Table 6 is obtained. As a second estimate of $\sigma^2_{\rho\tau}$ one could obtain the expected value of E'_{vv} in Table 6 for fixed w_{ij} . Then, Table 6 could be recomputed using the second estimate of $\sigma^2_{\rho\tau}$ and $\hat{\sigma}^2_\epsilon$. Table 6 could be recomputed and a third estimate of $\sigma^2_{\rho\tau}$ obtained. The process could be repeated until estimates of $\sigma^2_{\rho\tau}$ stabilize.

For a three-way or higher-way classification Henderson's [11] Method 1 probably should be used to obtain the estimated variance components if the degrees of freedom are fairly large, say greater than 20 to 30, for each sum of squares considered. Also, the three-way classification could be collapsed into a two-way classification for certain experimental situations in order to simplify the analysis. In still other situations, a Case II analysis might be necessary for two factors whereas the third factor and interactions with the third factor would involve a Case III analysis.

NUMERICAL EXAMPLES*

An experiment was conducted on the effect of different humidity treatments planted at different times (6/15, 7/15, 8/15, and 9/15) on roses over a period of months. One characteristic measured was the number of salable roses per month. The data for Y = number of salable roses in Tables 7 and 9 represent a selected sample of data from the total experiment. The X covariate represents the position on the greenhouse bench and is used to control variation within a replicate [see 7, sec. XVI 11]. The two replicates were on opposite greenhouse benches. For the data in Table 7, the positions were grouped into 5 sets of 3 each and given the numbers 1 to 5. A Case I analysis of covariance for these data and the treatment means adjusted for regression are presented in Table 8. Since this is a 2×5 table, Snedecor's [16, sec. 11.11] analysis of variance procedure may be followed for the Y and X variates, and for the cross products, or the entire analysis follows from a direct application of the formulae given for a Case I analysis.

TABLE 7

NUMBER OF SALABLE FLOWERS (ROSES), Y_{ijk}, OPENED IN ONE MONTH (DECEMBER) AND LOCATION ON GREENHOUSE BENCH, X_{ijk}—6/15 PLANTING

	Rep. I					Rep. II					Total		
Treatment	n_{i1}	Y	X	Y	X	n_{i2}	Y	X	Y	X	$n_{i.}$	$Y_{i..}$	$X_{i..}$
1	2	27	5	18	5	1	19	4	—	—	3	64	14
2	1	31	3	—	—	2	52	5	57	5	3	140	13
3	1	34	2	—	—	2	52	1	33	1	3	119	4
4	1	38	4	—	—	2	60	2	45	2	3	143	8
5	1	31	1	—	—	2	50	3	40	3	3	121	7
Total	6	$Y_{.1.}=179$; $X_{.1.}=20$				9	$Y_{.2.}=408$; $X_{.2.}=26$				15	587	46

The data in Table 9 are the number of salable roses obtained for the month of April from the 7/15 planting, Y, and the position on the greenhouse bench, X. The 15 positions were not grouped as they were in Table 7, but represent the location of the treatment in one of the 15 possible positions. It is assumed, as before, that there was a linear

*The data for this example were obtained through the courtesy of Dr. R. W. Langhans, Dept. of Floriculture, Cornell University, from an experiment reported in his doctoral dissertation.

TABLE 8

CASE I COVARIANCE ANALYSIS FOR DATA IN TABLE 7.

Source of variation	D.f.	Sums of products y^2	xy	x^2
Total (uncorrected)	15	25407	1755	174
Correction for mean	1	22971.27	1800.133	141.067
Total	14	2435.73	−45.133	32.933
Replicate (ign. treat.; elim. mean)	1	864.90	−24.800	0.711
Treatment (elim. mean and rep.)	4	912.70	−47.667	22.889
Residual	9	658.13	27.334	9.333

	D.f.	Adj. sum of squares	Mean square
Residual (elim. reg., treat., rep. and mean)	8	578.08	72.26
Treat. + residual	12	1558.00	—
Treatment (elim. reg., rep., and mean)	4	979.92	244.98

$\hat{\mu} + \hat{\tau}_1 = 18.55$ $\hat{\mu} + \hat{\tau}_3 = 42.85$ $\hat{\mu} + \hat{\tau}_5 = 40.58$
$\hat{\mu} + \hat{\tau}_2 = 41.06$ $\hat{\mu} + \hat{\tau}_4 = 46.94$ $\hat{\rho}_1 = -5.7 = -\hat{\rho}_2$

gradient from one end of the bench to the other. The Case II covariance analysis for the data of Table 9 is presented in Table 10, and is obtained by a direct application of the formulae given for a Case II analysis. The adjusted treatment means are 80.78, 74.15, 81.02, 63.78 and 61.78 for treatments 1, 2, 3, 4, and 5 respectively.

TABLE 9

NUMBER OF SALABLE FLOWERS (ROSES), Y_{ijh}, OPENED IN ONE MONTH (APRIL) AND LOCATION ON GREENHOUSE BENCH, X_{ijh}—7/15 PLANTING.

Treatment	Rep. I					Rep. II					Total		
	n_{i1}	Y	X	Y	X	n_{i2}	Y	X	Y	X	$n_{i.}$	$Y_{i..}$	$X_{i..}$
1	1	102	15	—	—	2	71	10	79	11	3	252	36
2	2	84	9	81	7	1	76	14	—	—	3	241	30
3	2	67	5	83	4	1	74	2	—	—	3	224	11
4	1	71	11	—	—	2	51	4	63	5	3	185	20
5	1	53	2	—	—	2	63	8	61	7	3	177	17
Total	7	$Y_{.1.}=541$; $X_{.1.}=53$				8	$Y_{.2.}=538$; $X_{.2.}=61$				15	1079	114

TABLE 10

CASE II ANALYSIS OF DATA IN TABLE 9.

Source of variation	D.f.	Sums of products y^2	xy	x^2
Total (elim. mean)	14	2426.93	447.60	229.60
Replicate (ign. tr. and inter.; elim. mean)	1	376.00	−2.01	0.01
Treatment (elim. rep. and mean; ign. inter.)	4	1320.97	303.81	136.12
I = Interaction (elim. tr., rep., and mean)	4	491.46	139.80	89.47
S = With subclasses	5	238.50	6.00	4.00
T = Treatment (elim. rep., inter., and mean)	4	1606.67	367.33	172.53

	D.f.	Adj. sum of squares	Mean square
S (adjusted for regression)	4	229.50	57.38
$S + I$ (" " ")	8	502.53	—
I (" " ")	4	273.03	68.26
$S + T$ (" " ")	8	1055.64	—
T (" " ")	4	826.14	206.54

The data from Table 9 are used to illustrate a Case III linear covariance analysis. The computational form for the analysis is given in Table 11. The first set of weights are obtained from the variance components estimated from the data (Tables 4 and 10) [see 8]:

$$\hat{\sigma}_\epsilon^2 = 57.38;$$

$$\hat{\sigma}_{\rho\tau}^2 = \frac{1}{k_0}\left\{\frac{I'_{yy}}{(r-1)(v-1)} - \frac{S'_{yy}}{n_{..} - rv - 1}\right\}$$

$$= \frac{1}{1.014}\{68.26 - 57.38\} = 10.73,$$

where

$$k_0 = \frac{1}{r_{\cdot} - r - v + 1}\left\{n_{..} + \sum_i \sum_j n_{ij}^2\left(\frac{1}{n_{..}} - \frac{1}{n_{i\cdot}} - \frac{1}{n_{\cdot j}}\right) - \frac{I_{xx}}{I_{xx} + S_{xx}}\right\}$$

$$= \frac{1}{4}\left(15 + \frac{25}{15} - \frac{25}{3} - \frac{11}{7} - \frac{14}{8} - \frac{89.47}{89.47 + 4.00}\right) = 1.014.$$

Since $n_{ij} = 1$ or 2 the two weights are (formula (III–4))

$$w_{11} = \frac{1}{10.73 + 57.38} = .015$$

and

$$w_{12} = \frac{2}{2(10.73) + 57.38} = .025.$$

Since there are only two weights, the coded weights $w_{11} = 3$ and $w_{12} = 5$ will be used to simplify the computations. The actual weights may be simpler to use than coded weights in examples where the n_{ij} vary considerably.

With the computed weights Table 11 can now be completed. Applying the formulae given for a Case III analysis, Table 12 is constructed. The various estimates used to obtain the sums of products are the μ' , τ_i' , ρ_j' , μ_x' , τ_{xi}' , and ρ_{xj}' values in Table 12.

A second estimate of $\sigma_{\rho\tau}^2$ is obtained from Tables 6 and 12 as

$$\acute{\sigma}_{\rho\tau}^2 = \frac{1}{1.014}\left\{\frac{767.86}{27.5 + 2.6875 - 13.4375 - 5.3797} - 57.38\right\} = 10.01,$$

where k_0 is computed in the same manner as described above and where

$$27.5 + 2.6875 + 13.4375 - 5.3791$$
$$= \sum\sum \frac{w_{ij}}{n_{ij}} + \frac{\sum\sum w_{ij}}{n_{ij}}\left(\frac{1}{w_{..}} - \frac{1}{w_{i.}} - \frac{1}{w_{.j}}\right).$$

Since the second estimate of $\sigma_{\rho\tau}^2$ is almost identical with the first estimate of $\sigma_{\rho\tau}^2$, which is $\acute{\sigma}_{\rho\tau}^2 = 10.73$, the second set of weights would be almost identical with the first. Likewise, the new analysis of variance table and the adjusted effect differences would be very similar. Therefore, the iterative analysis of variance stops here and the adjusted treatment means are obtained (Table 12). Under the assumption of fixed weights the tests of significance suggested previously may be carried out using the interaction mean square as the error mean square.

After obtaining $\hat{\beta}_1 = 393.19/251.63 = 1.562572$, the $\hat{\mu} + \hat{\tau}_i$ and $\hat{\rho}_j$ values may be computed. The two equations for the $\hat{\rho}_j$'s are (III–12):

$$\frac{75}{8}\hat{\rho}_1 - \frac{75}{8}\hat{\rho}_2 = 1465.5 - \tfrac{3}{8}(681.0 + 498.0 + 469.0)$$
$$- \tfrac{5}{8}(640.5 + 597.0)$$
$$- 1.562572[146.5 - \tfrac{3}{8}(97.5 + 55.5 + 43.5)$$
$$- \tfrac{5}{8}(82.0 + 28.5)]$$
$$= 74.0625 - 5.859645 = 68.202855;$$
$$\hat{\rho}_1 = -\hat{\rho}_2 .$$

TABLE 11

COMPUTATIONAL FORM FOR A CASE III COVARIANCE ANALYSIS FOR THE DATA OF TABLE 9.

Treatment	Replicate I		Replicate II		Sums		Means	
							$\bar{y}$	$\bar{x}$
1	$\bar{y}_{11.} = 102.0$	$\bar{x}_{11.} = 15.0$	$\bar{y}_{12.} = 75.0$	$\bar{x}_{12.} = 10.5$	177.0	25.5		
	$w_{11} = 3$	$n_{11} = 1$	$w_{12} = 5$	$n_{12} = 2$	8	3		
	$w_{11}\bar{y}_{11.} = 306.0$	$w_{11}\bar{x}_{11.} = 45.0$	$w_{12}\bar{y}_{12.} = 375.0$	$w_{12}\bar{x}_{12.} = 52.5$	681.0	97.5	85.1250	12.1875
2	$\bar{y}_{21.} = 82.5$	$\bar{x}_{21.} = 8.0$	$\bar{y}_{22.} = 76.0$	$\bar{x}_{22.} = 14.0$	158.5	22.0		
	$w_{21} = 5$	$n_{21} = 2$	$w_{22} = 3$	$n_{22} = 1$	8	3		
	$w_{21}\bar{y}_{21.} = 412.5$	$w_{21}\bar{x}_{21.} = 40.0$	$w_{22}\bar{y}_{22.} = 228.0$	$w_{22}\bar{x}_{22.} = 42.0$	640.5	82.0	80.0625	10.2500
3	$\bar{y}_{31.} = 75.0$	$\bar{x}_{31.} = 4.5$	$\bar{y}_{32.} = 74.0$	$\bar{x}_{32.} = 2.0$	149.0	6.5		
	$w_{31} = 5$	$n_{31} = 2$	$w_{32} = 3$	$n_{32} = 1$	8	3		
	$w_{31}\bar{y}_{31.} = 375.0$	$w_{31}\bar{x}_{31.} = 22.5$	$w_{32}\bar{y}_{32.} = 222.0$	$w_{32}\bar{x}_{32.} = 6.0$	597.0	28.5	74.6250	3.5625
4	$\bar{y}_{41.} = 71.0$	$\bar{x}_{41.} = 11.0$	$\bar{y}_{42.} = 57.0$	$\bar{x}_{42.} = 4.5$	128.0	15.5		
	$w_{41} = 3$	$n_{41} = 1$	$w_{42} = 5$	$n_{42} = 2$	8	3		
	$w_{41}\bar{y}_{41.} = 213.0$	$w_{41}\bar{x}_{41.} = 33.0$	$w_{42}\bar{y}_{42.} = 285.0$	$w_{42}\bar{x}_{42.} = 22.5$	498.0	55.5	62.2500	6.9375
5	$\bar{y}_{51.} = 53.0$	$\bar{x}_{51.} = 2.0$	$\bar{y}_{52.} = 62.0$	$\bar{x}_{52.} = 7.5$	115.0	9.5		
	$w_{51} = 3$	$n_{51} = 1$	$w_{52} = 5$	$n_{52} = 2$	8	3		
	$w_{51}\bar{y}_{51.} = 159.0$	$w_{51}\bar{x}_{51.} = 6.0$	$w_{52}\bar{y}_{52.} = 310.0$	$w_{52}\bar{x}_{52.} = 37.5$	469.0	43.5	58.6250	5.4375
Sums	383.5	40.5	344.0	38.5	727.5	79.0		
	19	7	21	8	40	15		
	1465.5	146.5	1420.0	160.5	2885.5	307.0	72.1375	7.6750

TABLE 12

COVARIANCE ANALYSIS FOR THE DATA IN TABLE 11.

Source of variation	D.f.	Sum of products y^2	xy	x^2
Total (on means)	9	6111.49	1273.29	648.78
Replicate (ign. treat.)	1	902.62	6.42	.05
Treatment (elim. rep.) = T	4	3826.62	873.68	397.10
Error = E	4	1382.25	393.19	251.63
Within subclasses = S	5	238.50	6.00	4.00
Treatment (ign. rep.)	4	4144.15	850.48	395.65

Source of variation	Regression D.f.	Regression Ss	Adjusted sum of squares D.f.	Adjusted sum of squares Ss	Adjusted sum of squares Ms
S	1	9.00	4	229.50	57.38
E	1	614.39	3	767.86	255.95
$E + T$	1	2474.00	7	2734.87	—
Treatment (elim. other effects)			4	1967.01	491.75

$\mu' = 72.3350$ $\quad \tau_4' = -9.0975$ $\quad \tau_{x1}' = 4.5525$ $\quad \tau_{x5}' = -2.1975$
$\tau_1' = 13.7775$ $\quad \tau_5' = -12.7225$ $\quad \tau_{x2}' = 2.5150$ $\quad \mu_x' = 7.6850$
$\tau_2' = 6.7400$ $\quad \rho_1' = 3.9500$ $\quad \tau_{x3}' = -4.1725$ $\quad \rho_{x1}' = 0.2000$
$\tau_3' = 1.3025$ $\quad \rho_2' = -3.9500$ $\quad \tau_{x4}' = -0.6975$ $\quad \rho_{x2}' = -0.2000$

$\hat{\beta} = 1.5000$ $\quad \hat{\mu} + \hat{\tau}_3 = 80.0245$
$\hat{\beta}_1 = 1.5626$ $\quad \hat{\mu} + \hat{\tau}_4 = 64.1946$
$\hat{\mu} + \hat{\tau}_1 = 78.8661$ $\quad \hat{\mu} + \hat{\tau}_5 = 62.9134$
$\hat{\mu} + \hat{\tau}_2 = 75.0123$ $\quad \hat{\rho}_1 = 3.6375 = -\hat{\rho}_2.$

Alternatively, the two equations involving $\hat{\rho}_1$ and $\hat{\rho}_2$ from formula (III–9) are:

$$\hat{\rho}_1\left[\frac{75}{8} - \frac{1}{253.13}\left(146.5 - \tfrac{3}{8}(196.5) - \tfrac{5}{8}(110.5)\right)^2\right] - \hat{\rho}_2\left(\frac{75}{8} + \frac{1}{253.13}(3.75)(-3.75)\right)$$

$$= 74.0625 - \frac{422.81}{253.13}(3.75) = 67.798772,$$

and

$$\hat{\rho}_1 = -\hat{\rho}_2 ,$$

where $253.13 = 648.78 - 395.65$ and $422.81 = 1273.29 - 850.48$.

SUMMARY

Variance and covariance analyses have been classified under three categories, *viz.*, Case I, interaction absent; Case II, interaction present and the effects assumed to be fixed effects; and Case III, interaction present and the interaction effects and at least one of the main effects of the factors represented in the interactions assumed to be random effects. The statistical procedures for the three cases have been derived for two-way and three-way classifications and are illustrated with numerical examples for the two-way classification with a covariate. The procedures for a q-way classification with b covariates are indicated.

REFERENCES

[1] Bartlett, M. S. [1936] A note on the analysis of covariance, *J. Agr. Sci. 26:* 488–491.

[2] Bartlett, M. S. [1937] Some examples of statistical methods of research in agriculture and applied biology, *J. Roy. Stat. Soc. Suppl. 4:* 137–183.

[3] Cochran, W. G. [1943] Analysis of variance for percentages based on unequal numbers, *J. Amer. Stat. Assoc. 38:* 287–301.

[4] Das, M. N. [1953] Analysis of covariance in two-way classification with disproportionate cell frequencies, *J. Indian Soc. Agric. Stat. 5:* 161–178.

[5] Day, B. and Fisher, R. A. [1937] The comparison of variability in populations having unequal means. An example of the analysis of covariance with multiple dependent and independent variates, *Annals of Eugenics 7:* 333–348.

[6] Federer, W. T. [1954] *Covariance analysis in a two-way classification with unequal numbers in the subclasses,* BU-51-M (mimeograph), Cornell University, Ithaca, New York, April.

[7] Federer, W. T. [1955] *Experimental design—theory and application,* Macmillan Company, New York.

[8] Federer, W. T. *Covariance analysis for unbalanced two-way classifications,* Cornell Univ. Agric. Expt. Sta. Memoir (accepted for publication).

[9] Hazel, L. N. [1946] The covariance analysis of multiple classification tables with unequal subclass numbers, *Biometrics 2:* 21–25.

[10] Henderson, C. R. [1948] *Estimation of general, specific, and maternal combining abilities in crosses among inbred lines of swine,* Ph.D. Thesis, Iowa State College Library, Ames, Iowa.

[11] Henderson, C. R. [1953] Estimation of variance and covariance components, *Biometrics 9:* 226–252.

[12] Nair, K. R. [1941] A note on the method of "fitting of constants" for analysis of non-orthogonal data arranged in a double classification, *Sankhyā 5:* 317–328.

[13] Outhwaite, A. D., and Rutherford, A. [1955] Covariance analysis as an alternative to stratification in the control of gradients, *Biometrics 11:* 431–440.

[14] Quenouille, M. H. [1948] The analysis of covariance and non-orthogonal comparisons, *Biometrics 4:* 240–246.
[15] Rao, C. R. [1946] On the linear combination of observations and the general theory of least squares, *Sankhyā 7:* 237–256.
[16] Snedecor, G. W. [1946] *Statistical methods*, 4th ed. Iowa State College Press, Ames, Iowa.
[17] Snedecor, G. W. and Cox, G. M. [1935] *Disproportionate subclass numbers in tables of multiple classification*, Iowa Agric. Expt. Sta. Res. Bul. 180: 233–272.
[18] Stevens, W. L. [1948] Statistical analysis of a non-orthogonal tri-factorial experiment, *Biometrika 35:* 346–367.
[19] Wilks, S. S. [1938] Analysis of variance and covariance in non-orthogonal data, *Metron 13:* 141–154.
[20] Yates, F. [1933] The principles of orthogonality and confounding in replicated experiments, *J. Agric. Sci. 23:* 108–145.
[21] Yates, F. [1934] The analysis of multiple classifications with unequal numbers in the different classes, *J. Amer. Stat. Assoc. 29:* 51–66.
[22] Yates, F. [1938] Orthogonal functions and tests of significance in the analysis of variance, *J. Roy Stat. Soc. Suppl. 5:* 177–180.

THE ANALYSIS OF COVARIANCE WITH INCOMPLETE DATA

G. N. Wilkinson
Commonwealth Scientific and Industrial Research Organization, Division of Mathematical Statistics, Adelaide, Australia

1. *Introduction*

In a paper [4] submitted for publication in this journal, the author has presented simplified methods for setting up and solving equations for missing values. The methods apply to those designs and data for which normal theory, with linear model, provides the appropriate analysis, and have been extended to cover covariance analyses.

A. T. James, in a personal communication, drew the author's attention to the fact that concomitant measurements corresponding to missing observations are irrelevant to the analysis of existing observations, and could therefore be replaced by hypothetical values more convenient for determining missing values and for the subsequent analysis of covariance on the completed data. Application of this suggestion has greatly simplified the author's results, and as the procedure is so simple, it has been thought worth while to outline it in this advance note.

Since deriving these results, the author's attention has been drawn to a paper by Barnard [1] in which the method of fitting values appears to be the same as that proposed here. Barnard's description is very brief, however, being incidental to the main content of the paper (analysis of a crop-weather scheme), and seems to have been overlooked in the literature and textbooks. The present paper gives details of the fitting process, and also of deriving standard errors and exact significance tests.

2. *Basic procedure*

Suppose that a covariance analysis is required on a set of observations of a variate y and the corresponding measurements of p concomitant variates $x_1 , x_2 , \cdots , x_p$.

If some observations on y are missing,

(i) discard all measurements of $x_1 , x_2 , \cdots , x_p$ that correspond to the missing observations on y,

(ii) fit a set of missing values for y, as though for an ordinary analysis (ignoring covariance),

(iii) with exactly parallel calculations fit sets of missing values for $x_1 , x_2 , \cdots , x_p$, to replace those discarded,

(iv) carry out the covariance analysis on the completed data for y and $x_1 , x_2 , \cdots , x_p$.

These steps provide the simplest set of calculations for exact analysis, and formal justification of the procedure is given in [4].

3. *Estimation of the missing values*

Each set of missing values could, of course, be determined iteratively, but solution of missing value equations is much the shorter method. Each set of equations has the form

$$\mathbf{A}\boldsymbol{u} = \boldsymbol{E}$$

where $\mathbf{A}$ is the matrix of coefficients, $\boldsymbol{u}$ is the vector of missing value unknowns, and $\boldsymbol{E}$ is a vector of initial plot estimates for those plots with missing observations. The initial plot estimates are determined by substituting zeros for the missing (or discarded values) in the data, that is, by using incomplete totals, etc., as though they were complete.

The matrix $\mathbf{A}$ is determined by the configuration in the design of the plots with missing values, and is the same for all $p + 1$ sets of equations. Inversion of this one matrix therefore provides estimates of missing values for all the variates $y, x_1, x_2, \cdots, x_p$. Rapid methods for forming and inverting the matrix of missing value equations are given in [4]. Each element of the matrix corresponds to a pair of missing values (the diagonal elements corresponding to the identity pairs (u, u)), and is determined by their relationship in the design, for example, "Same Block" or "Identical". The author has prepared for a number of the standard designs, tables which give the various possible relationships for each design, and the corresponding elements from which the matrix may be constructed. One such table (Table 4) is given.

4. *Numerical example*

So that comparison may be made with the parallel analysis on a complete set of observations, the data used by Quenouille [3] (pp. 40–47) have been rendered incomplete by removing three observations at random.

The data are the results of an insulin-response experiment, with four dosage levels (treatments) A, B, C and D, carried out on eight rabbits on each of four occasions (phases), according to the Latin square design shown in Table 1. The percentage falls in blood sugar (y) are given in Table 2, and the initial blood sugars (x) in Table 3. The three measurements of x corresponding to the missing observations on y are omitted in Table 3, in accordance with 2(i) above. Incomplete totals are shown in separate columns, and completed totals, obtained after estimating missing values, are indicated by heavy type.

Calculations for determining the missing values u, v and w are

TABLE 1

DESIGN OF INSULIN-RESPONSE EXPERIMENT

Phase	Rabbit 1	2	3	4	5	6	7	8
1	*C*	*A*	*B*	*D*	*C*	*A*	*D*	*B*
2	*B*	*D*	*C*	*A*	*D*	*B*	*C*	*A*
3	*A*	*C*	*D*	*B*	*A*	*C*	*B*	*D*
4	*D*	*B*	*A*	*C*	*B*	*D*	*A*	*C*

TABLE 2

PERCENTAGE FALL IN BLOOD SUGAR (y)

Phase	Rabbit 1	2	3	4	5	6	7	8	Totals Incomplete	Complete
1	32.7	11.2	23.2	48.1	35.1	27.2	36.0	40.0		253.5
2	26.2	31.8	28.9	18.7	37.2	$w(B)$	37.8	25.8	206.4	***242.3***
3	−4.0	14.0	27.5	$v(B)$	2.8	36.2	28.4	50.7	155.6	***179.7***
4	$u(D)$	16.5	21.2	40.2	12.7	47.7	25.1	39.4	202.8	***236.7***
Incomplete total	54.9			107.0		111.1			818.3	
Complete total	***88.8***	73.5	100.8	***131.1***	87.8	***147.0***	127.3	155.9		***912.2***

Dosage	*A*	*B*	*C*	*D*
Incomplete total		147.0		279.0
Complete total	128.0	***207.0***	264.3	***312.9***

shown in Table 5. For computational convenience the equations have been multiplied through by a factor of 16. The matrix **A** may be determined from first principles, or more rapidly from Table 4: the diagonal elements in 16**A** are 9; the missing values (u, v), (u, w) are in different rows and columns, and have different treatments, so that the corresponding elements in the matrix are $+1$; the pair (v, w), in different rows and columns, have the same treatments, hence the corresponding element is -1.

For a 3×3 matrix the simplest inversion procedure is to form the

TABLE 3

Initial Blood Sugar (x)

Phase	Rabbit 1	2	3	4	5	6	7	8	Totals Incomplete	Complete
1	107	91	99	93	103	88	94	89		764
2	94	93	90	77	94	$w(B)$	91	89	628	***707.9***
3	75	83	91	$v(B)$	87	79	81	84	580	***653.9***
4	$u(D)$	86	97	87	101	81	87	95	634	***730.9***
Incomplete total	276			257		248			2606	
Complete total	***372.9***	353	377	***330.9***	385	***327.9***	353	357		***2856.7***

Dosage	A	B	C	D
Incomplete total		550		630
Complete total	691	***703.8***	735	***726.9***

TABLE 4

Elements of A-matrix for Missing Value Equations. Row-coupled Latin Squares

(u, v) relation	$r\,p \times p$ squares $a_{uv} \times rp^2$	2 4 × 4 squares $a_{uv} \times 16$
Identical	$(rp - 2)(p - 1)$	9
Same C	$-(rp - 2)$	−3
Same R, same T	$-2(p - 1)$	−3
" diff. T	$-(p - 2)$	−1
Diff. R and C, same T	$-(p - 2)$	−1
" diff. T	2	1

matrix of cofactors $[A_{ij}]$, and divide by the determinant $|\,\mathbf{A}\,|$, which may be calculated by taking the sum of products of any two corresponding rows in $\mathbf{A}$ and $[A_{ij}]$.

The formula for a plot estimate in this design is, in the usual notation,

$$E_{ijk} = (2R_i + 4C_j + 2T_k - G)/16.$$

TABLE 5

ESTIMATION OF THE MISSING VALUES

$\mathfrak{y}$	16 A			$[A_{ij}] \times (16)^2$			Initial plot estimates $16E(y)$	$16E(x)$	Missing values $\mathfrak{y}(y)$	$\mathfrak{y}(x)$
u	9	1	1	80	−10	−10	364.9	1026	33.9	96.9
v	1	9	−1	−10	80	10	214.9	682	24.1	73.9
w	1	−1	9	−10	10	80	332.9	742	35.9	79.9
				$\lvert\mathbf{A}\rvert \times (16)^3 = 700$						

TABLE 6

ANALYSIS OF VARIANCE AND COVARIANCE

Source of variation	D.f.	S.s. (y)	S.p.	S.s. (x)
Rabbits	7	1612.47	−674.02	749.19
Phases	3	407.92	529.70	804.33
Dosages	3	2370.84	546.46	155.04
Residual	15	519.79	106.37	297.30
Total	28	4911.02	508.51	2005.86

ANALYSIS OF RESIDUAL VARIANCE

Source of variation	D.f.	S.s. (y)	M.s.
Regression on x	1	38.06	38.06
Residual (elim. regression)	14	481.73	34.41
Residual (ignoring regression)	15	519.79	34.65

Treatment means	A	B	C	D	Average
Initial blood sugar (x)	86.3750	87.9750	91.8750	90.8625	89.2719
Deviations	−2.8969	−1.2969	2.6031	1.5906	0
Fall in blood sugar (y)					
Unadjusted	16.0000	25.8750	33.0375	39.1125	28.5063
Adjusted to $\bar{x} = 89.2719$ ($b = 0.35781$)	17.04	26.34	32.11	38.54	

Using the incomplete totals of Tables 2 and 3, two sets of three initial plot estimates $E(y)$ and $E(x)$ (multiplied by 16), are calculated. For example,

$$16\,E_v(y) = 2 \times 155.6 + 4 \times 107.0 + 2 \times 147.0 - 818.3 = 214.9.$$

The missing values are now determined by taking the sums of products $\mathbf{A}^{-1}\boldsymbol{E}$. For example, in Table 5,

$$w(x) = (-1 \times 1026 + 1 \times 682 + 8 \times 742)/70 = 79.9.$$

The analysis of variance and covariance on the completed data is shown in Table 6, with the analysis of residual variance and a table of unadjusted and adjusted treatment means.

5. *Variance-covariance matrix of adjusted treatment means*

If any treatment comparison involves fitted missing values, its usual variance must be increased somewhat. The adjustment to the variance is easily computed from the inverse of the matrix in the missing value equations.

Suppose the treatment comparison £ contains $l_1u + l_2v + l_3w, = \boldsymbol{l'u}$, say. The increase in variance of £ due to fitted missing values is

$$(l'\mathbf{A}^{-1}l)\sigma^2.$$

Similarly, the adjustment to the covariance of two treatment comparisons £ and $\mathfrak{M}$ is

$$(l'\mathbf{A}^{-1}m)\sigma^2.$$

Hence in our example (see Table 7) the variance-covariance matrix of unadjusted treatment means is easily obtained. $\bar{y}_A$ and $\bar{y}_C$ do not involve missing values, so their variance is $\sigma^2/8$. $\bar{y}_B$ contains $(v + w)/8$. The increase in variance is therefore

$$\frac{16}{70} \times [0 \quad \tfrac{1}{8} \quad \tfrac{1}{8}] \begin{bmatrix} 8 & -1 & -1 \\ -1 & 8 & 1 \\ -1 & 1 & 8 \end{bmatrix} \begin{bmatrix} 0 \\ \tfrac{1}{8} \\ \tfrac{1}{8} \end{bmatrix} \times \sigma^2 = \frac{9}{140}\sigma^2 = 0.0643\sigma^2.$$

Similarly, the increase in variance of $\bar{y}_D$ is $\sigma^2/35 = 0.0286\sigma^2$. The only covariance increase is for $\bar{y}_B$ and $\bar{y}_D$, and is

$$\frac{16}{70} \times [0 \quad \tfrac{1}{8} \quad \tfrac{1}{8}] \begin{bmatrix} 8 & -1 & -1 \\ -1 & 8 & 1 \\ -1 & 1 & 8 \end{bmatrix} \begin{bmatrix} \tfrac{1}{8} \\ 0 \\ 0 \end{bmatrix} \times \sigma^2 = -\frac{1}{140}\sigma^2 = -0.0071\sigma^2.$$

TABLE 7

VARIANCE-COVARIANCE MATRIX OF TREATMENT MEANS ADJUSTED TO THE MEAN INITIAL BLOOD SUGAR, 89.2719 (MULTIPLY ENTRIES BY σ^2)

Unadjusted

A	0.125	0	0	0
B	0	0.125	0	0
C	0	0	0.125	0
D	0	0	0	0.125

Adjustment for missing values

0	0	0	0
0	0.064	0	−0.007
0	0	0	0
0	−0.007	0	0.029

+

Adjustment for covariant adjustments

0.028	0.013	−0.025	−0.016
0.013	0.006	−0.011	−0.007
−0.025	−0.011	0.023	0.014
−0.016	−0.007	0.014	0.009

=

0.153	0.013	−0.025	−0.016
0.013	0.195	−0.011	−0.014
−0.025	−0.011	0.148	0.014
−0.016	−0.014	0.014	0.162

If the treatment means $\bar{y}$ are adjusted to a given value x_0 of a concomitant variate, that is, to $\bar{y} - b(\bar{x} - x_0)$, the corresponding adjustment to the variance, when there are no missing values is

$$(\bar{x} - x_0)^2 \operatorname{Var}(b) = \frac{(\bar{x} - x_0)^2}{\text{Residual S.s. for } x} \sigma^2.$$

Likewise, the adjustment to the covariance of two means is

$$\frac{(\bar{x}_1 - x_0)(\bar{x}_2 - x_0)}{\text{Residual S.s. for } x} \sigma^2.$$

The generalization for covariant adjustment on one or more of several variates is straightforward. These adjustments to variance and covariance also apply, without modification, when missing values are fitted for some observations y, provided that corresponding x-values have been fitted as indicated above, for the variance of b is then correctly determined by the residual sum of squares for x in the analysis of covariance for the completed data.

The variance-covariance matrix of adjusted treatment means is

therefore the sum of three components:

(i) the variance-covariance matrix of unadjusted means (ignoring missing values),

(ii) the adjustment to (i) for missing values,

(iii) the adjustment to (i) for the covariant adjustments.

These components are set out in Table 7. It will be seen that treatment B, for which there are two missing values, suffers an appreciable increase in variance.

6. *Significance tests*

Yates [6] has shown that the usual analysis of variance on data with fitted missing values yields a slightly exaggerated treatment sum of squares. In principle, to obtain an exact test of significance, an auxiliary set of missing values should be fitted for an auxiliary analysis of variance which ignores treatments. The correct treatment sum of squares is then the difference between the residual sum of squares (ignoring treatments) and the residual sum of squares (eliminating treatments).

In practice it is not necessary to fit another set of missing values. Yates [6] and Cornish [2] have provided formulae for correcting the treatment sum of squares in randomized block and incomplete block experiments, and the author [5] has shown that the general form of such corrections is

$$\eta' \mathbf{A}_{\text{aux}}^{-1} \eta,$$

where $\mathbf{A}_{\text{aux}}$ is the matrix of the missing value equations for auxiliary estimates, and η is the vector of residuals (ignoring treatments) of the fitted values (main estimates). In an orthogonal design, the residual (ignoring treatment) of a fitted value is just the estimated treatment effect for that plot.

The extension of this argument to analysis of covariance is fairly obvious. Residual vectors (ignoring treatments) of fitted missing values are obtained for all variates. Denote these vectors by η, ξ_1 , ξ_2 , $\cdots$ ξ_p . The matrix of auxiliary missing value equations is the same for all variates. The "treatments" line of the analysis of variance and covariance is then corrected by subtracting the appropriate bilinear forms in η, ξ_1 , ξ_2 , $\cdots$ ξ_p . Thus the treatment sum of squares for y is reduced by $\eta' \mathbf{A}_{\text{aux}}^{-1} \eta$, the sum of products for x_1 and y by $\eta' \mathbf{A}_{\text{aux}}^{-1} \xi_1$, and the sum of squares for x_1 by $\xi_1' \mathbf{A}_{\text{aux}}^{-1} \xi_1$, etc. The significance test for treatments (eliminating covariance) is then constructed in the usual way from the analysis of variance and covariance.

The experimental design ignoring treatments, for the data given in section 4, reduces to a two-way classification, phases $\times$ rabbits. As all

missing values are in different rows and columns, the matrix for the auxiliary missing value equations may be shown to be

$$\mathbf{A}_{\text{aux}} = \frac{1}{32}\begin{vmatrix} 21 & 1 & 1 \\ 1 & 21 & 1 \\ 1 & 1 & 21 \end{vmatrix} = \frac{20}{32}\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} + \frac{1}{32}\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix},$$

the inverse of which

$$\mathbf{A}_{\text{aux}}^{-1} = \frac{32}{20}\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} - \frac{32}{460}\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix}.$$

This gives the correction formulae $\eta'\mathbf{A}_{\text{aux}}^{-1}\eta = (32/20)\sum \eta^2 - (32/460)(\sum \eta)^2$, etc. The formula for the residual (ignoring treatments) of a fitted value u is, in the usual notation,

$$\eta_u = \frac{32u - 4R_u - 8C_u + G}{32} = \frac{4T_u - G}{32},$$

the relevant totals being those for the completed data. It is more convenient for computation to calculate the quantities (32η). These are given in Table 8. The correction formulae are then

$$\eta'\mathbf{A}_{\text{aux}}^{-1}\eta = \frac{1}{640}\sum(32\eta)^2 - \frac{1}{14720}(\sum 32\eta)^2, \quad \text{etc.}$$

TABLE 8

TREATMENT EFFECTS (UNADJUSTED) FOR u, v AND w ($\times$ 32)

Treatment	u (D)	v (B)	w (B)	Total
(32η)	339.4	−84.2	−84.2	171.0
(32ξ)	50.9	−41.5	−41.5	−32.1

The corrected analysis of covariance and the final analysis for treatment effects (eliminating covariance) are given in Table 9.

It should be pointed out here that the analyses in this paper have been performed for purposes of illustration only. It is apparent from the analysis of residual variance in Table 6, that in this instance no advantage is gained by adjusting for the concomitant variate in the analysis.

TABLE 9

CORRECTION OF TREATMENT SUMS OF SQUARES AND SUM OF PRODUCTS

Source of variation	D.f.	S.s. (y)	S.p.	S.s. (x)
Treatments (uncorrected)	3	2370.84	546.46	155.04
$-\sum(32\eta)^2/640$, etc.		−202.14	−37.91	−9.43
$+(\sum 32\eta)^2/14720$, etc.		1.99	−0.37	0.07
Treatments (corrected)	3	2170.69	508.18	145.68
Residual (eliminating treatments)	15	519.79	106.37	297.30
Residual (ignoring treatments)	18	2690.48	614.55	442.98

ANALYSIS FOR TREATMENTS (ELIMINATING REGRESSION ON x)

Source of variation	D.f.	S.s.	M.s.	F
Treatments (uncorrected)	3	1466.72	488.91	14.21
Treatments (exact)		1356.18	452.06	13.14
Residual (eliminating treatments)	14	481.73	34.41	
Residual (ignoring treatments)	17	1948.45		
		1837.91		

REFERENCES

[1] Barnard, M. M. [1936] An examination of the sampling observations on wheat of the crop-weather scheme. *J. Agric. Sc. 26:* 456–487.

[2] Cornish, E. A. [1940] The estimation of missing values in incomplete randomized block experiments. *Ann. Eug. 10:* 112–118.

Cornish, E. A. [1940] The estimation of missing values in quasi-factorial designs. *Ann. Eug. 10:* 137–143.

[3] Quenouille, M. H. [1953] *The design and analysis of experiment.* Charles Griffin and Company, Ltd., London.

[4] Wilkinson, G. N. Estimation of missing values for the analysis of incomplete data. Submitted for publication to *Biometrics.*

[5] Wilkinson, G. N. The analysis of variance and derivation of standard errors for incomplete data. Submitted for publication to *Biometrics.*

[6] Yates, F. [1933] The analysis of replicated experiments when the field results are incomplete. *Emp. J. Exp. Agric. 1:* 129–42.

STRATIFICATION, BALANCE, AND COVARIANCE

D. J. Finney
Department of Statistics and A.R.C. Unit of Statistics, University of Aberdeen, Aberdeen, Scotland

1. INTRODUCTION

Some experimenters seek to avoid the need for randomized block designs by balancing the experimental units assigned to different treatments in respect of an initial quality. A particularly common instance is that of an animal nutrition experiment in which the initial mean weights of animals on the different diets are deliberately equalized by suitable allocation of animals to treatments. Apart from difficulties of maintaining an objective allocation, this procedure may achieve almost the same end as does a randomized block design, in increasing the true precision of treatment comparisons by eliminating variation associated with the factor used in balancing. What is often forgotten is that, if the statistical analysis appropriate to a completely randomized design is used, the experimental error will be overestimated; the more successful the balancing is in increasing true precision, the more seriously will the estimated precision so calculated fall below the truth.

The explanation, and the manner in which the difficulty may be overcome, can be seen from comparison of four alternative experimental designs. Discussion of them will throw light on the practical utility of the balancing procedure.

2. THE FOUR DESIGNS

Suppose that t treatments are to be compared by an experiment on N animals, r being assigned to each treatment ($N = rt$). Results are to be assessed in terms of a final measurement, y, on each animal. The only information on the animals available at the start of the experiment is an initial (pre-treatment) measurement x, which is thought likely to show an approximately linear correlation (positive or negative) with y. Often x and y are measurements of the same kind, such as initial and final body weights, but this is not essential to the statistical argument.

Four different experimental designs are in common use:

I *Completely randomized*: Allot r animals to each treatment entirely at random.

II *Randomized block*: Stratify the available animals in blocks of t in such a way that each block is reasonably

homogeneous in respect of x. In each of r such blocks, select one animal at random for each treatment.

III *Simply balanced for treatments*: Proceed similarly to I, but modify the results of the randomization so as to make the mean values of x for the several treatments approximately equal.

IV *Randomized block, balanced for treatments*: Proceed similarly to II, but restrict random selection within the blocks so as to make the mean values of x for the several treatments approximately equal.

The process of allocating animals to their places in the experiment will vary a little according to whether there is a large, effectively infinite, stock from which to select or a fixed set of N animals all of which must be used. For design I, the difference is slight. For II, an infinite population would first be regarded as stratified into an infinite number of blocks of t, and a random selection of r blocks would be made for the experiment; on the other hand, if a particular N animals had to be used, they would presumably be stratified into the t largest values of x, the t next largest, and so on, and all blocks would be taken. If the population of available animals is infinite, one can conceive of a selection for III or IV being made at random from an infinite number of sets of N animals allocated so as to give exact equality of treatment means for x, though the mechanics of making such a selection are not easily visualized! If only N animals are available, there may be no allocation satisfying the condition exactly and, even if some arbitrary definition of approximate equality of means of x is adopted, there is not likely to be much scope for random selection amongst alternative allocations. Thus one immediate weakness in designs III and IV is the danger of bias or other invalidity entering into the statistical analyses through the method of allocation of animals to treatments. An objective rule of allocation that would at least come near to achieving balance could easily be based upon a preliminary arrangement of animals in descending order of x followed by assignment of the first t to treatments in one sequence, assignment of the next t in the reverse of this sequence, the next t in the original sequence, and so on. This, however, allows no scope for randomization except in the basic sequence of treatments. It is essentially the same as a systematic design for field-plot experiments that has been severely criticized by many statisticians over the last thirty years, and that is now generally discredited on account of the unpre-

dictable bias in estimation of experimental error that it may produce. To introduce a satisfactory randomness of allocation appears to be difficult, but this will be overlooked for the present; succeeding sections of the paper will show that, even if difficulties under this head can be removed, balancing has little to recommend it.

It is perhaps well to make clear at this point that the whole paper is written with reference to experiments in which the experimental unit or "plot" is a single animal, and measurements on different animals treated alike can be regarded as independent. Statisticians have frequently emphasized that, when "lots" of animals are penned together for uniform treatment (as is commonly done for chicks), the lot is to be regarded as the unit for analysis, and any calculation of an error variance from variation within lots is likely seriously to underestimate the truth. If lots were balanced so that all were nearly equal in mean or total value of an initial variate, x (whatever the variation of x within lots), the quantities $x_{.i}$, $y_{.i}$, defined in Section 3 should be taken to refer to totals or means of lots instead of to single animals. There would remain the possibilities of random allocation of these equalized lots to treatments (designs I, II) or a further balancing of any remaining differences between lots in respect of x, by allocation to treatments so that totals of x for all treatments are equal (designs III, IV again). The incentive to choose designs with this second balance over treatments ought to be much less when a first equalization of lots has been secured; if one is chosen, the objections that are to be raised below in respect of experiments on single animals will still apply.

A preliminary balancing of lots that is to be followed by a randomization according to design I or II falls quite outside the scope of this paper, and does not affect the formal validity of the statistical analysis for randomized experiments. However, another entirely different danger, perhaps manifesting itself as a breakdown of the model described in Section 3, needs to be considered when this type of balancing is adopted. An immediate consequence of the balancing is likely to be that some lots are much more varied than others in respect of the values of x possessed by their members, and effects of competition within lots may thereby be intensified. For example, if x represents initial weight in a feeding trial, the more variable lots might have quite different expectations of mean weight increase from those that were uniform, even the variance of a final measurement might be affected, and, in extreme instances, further trouble might arise from differential mortality. Of course, none of these disturbances will necessarily occur, but their possibility needs to be taken into account before such balance is introduced.

3. MODELS

Kempthorne and Wilk [1955] have described a general type of linear model for experimental observations that includes as one extreme the infinite population of available experimental units and as the other extreme an experiment limited to a particular set of units. This might be adapted to the present problem. However, since the chief interest lies not in comparing I with II but in comparing both of these with III and IV, and there is the great difficulty of defining III and IV exactly without invoking an infinite population, the discussion that follows will be presented in terms of an infinite model. The difference is not important in a qualitative sense; for I and II at least, the finite model can be reconciled with the infinite by suitable definition of symbols.

It will suffice to describe here the model and to define the symbols appropriate to II, leaving any modifications for other designs to be introduced as they are required. Provision must be made for studying the effects of analysing the experiment both without and with a covariance on the initial measurement. Although, for any one experiment in which the analysis of covariance has been applied, the precision of comparisons in respect of y would ordinarily be assessed subject to the constancy of certain functions of x, for the purpose of comparing the *average* precisions of different designs it is necessary to consider x as a random variate in the population from which animals are selected.

Write x_{ij} , y_{ij} for the initial and final measurements on the animal in block i that receives treatment j. Then the linear model for x can be written

$$x_{ij} = \xi + \theta_i + \alpha_{ij} , \tag{1}$$

where ξ is the general mean for the population, θ_i and α_{ij} are random variates all independent of one another and

$$E(\theta_i) = 0, \qquad E(\theta_i^2) = \Sigma_x^2 , \tag{2}$$

$$E(\alpha_{ij}) = 0, \qquad E(\alpha_{ij}^2) = \sigma_x^2 . \tag{3}$$

The model for y must make provision for a linear regression on the deviation of the corresponding x from its block mean. Hence

$$y_{ij} = \eta + \phi_i + \tau_j + \beta\alpha_{ij} + \epsilon_{ij} , \tag{4}$$

where η is the general mean, τ_j is the deviation from the general mean attributable to treatment j, β is the regression coefficient (here for simplicity assumed independent of treatment). The ϕ_i and ϵ_{ij} are random variates. The ϕ_i need not be independent of the θ_i , and in general there will be a regression for these block variates also, which will

be assumed linear. Writing

$$\phi_i = B\theta_i + \psi_i \ , \tag{5}$$

the model becomes

$$y_{ij} = \eta + \psi_i + \tau_j + B\theta_i + \beta\alpha_{ij} + \epsilon_{ij} \ . \tag{6}$$

In practice, B may be expected to be equal to β, both being the regression coefficient of y on x in the whole population which has been arbitrarily stratified on the basis of values of x. The more general model is retained because of the possibility that it may sometimes be appropriate. The random variates θ_i , ψ_i , α_{ij} , ϵ_{ij} are all independent of one another, and

$$E(\psi_i) = 0, \qquad E(\psi_i^2) = \Sigma_y^2 \ , \tag{7}$$

$$E(\epsilon_{ij}) = 0, \qquad E(\epsilon_{ij}^2) = \sigma_y^2 \ . \tag{8}$$

Note that Σ_y^2 , σ_y^2 are inter-block and intra-block variances for y after elimination of the appropriate regressions on x.

It is further convenient to write

$$\tau_1^2 + \tau_2^2 + \cdots + \tau_t^2 = (t - 1)T, \tag{9}$$

though T is not necessarily, or even commonly, a variance of a random variate. The symbols $\bar{y}_j$ and $\bar{y}_j'$ will be used for the mean values of y for treatment j without and with an adjustment for covariance on x calculated by the standard routine.

A further assumption, that all θ_i , α_{ij} , are normally distributed, will be introduced. This is not essential to the main argument, but it enables the average variance of treatment means adjusted by an analysis of covariance to be simply expressed; as shown elsewhere (Finney, 1946; an approximation for the final average is now replaced by the exact value), multiplication of the residual variance about the regression by a factor $[1 + 1/(\text{error d.f.} - 2)]$ gives a value for the effective variance per animal in comparisons between treatments. Without the assumption of normality, the corresponding expression must be written in terms of actual sums of squares of x for treatments and error, and is consequently more cumbrous, but qualitatively conclusions would be similar. No assumption that ψ_i , ϵ_{ij} are normally distributed need be made.

4. RANDOMIZED BLOCK DESIGN (DESIGN II)

Table 1 shows the expected mean squares and products in an analysis of variance and covariance. This follows a familiar pattern, except for the elaboration of separating the regression and residual components in the y analysis.

TABLE 1

EXPECTATIONS OF MEAN SQUARES AND PRODUCTS FOR RANDOMIZED BLOCK DESIGN (II)

Variation	D.f.	(x^2)	(xy)	(y^2)
Blocks	$r - 1$	$\sigma_x^2 + t\Sigma_x^2$	$\beta\sigma_x^2 + tB\Sigma_x^2$	$\beta^2\sigma_x^2 + \sigma_y^2 + t(B^2\Sigma_x^2 + \Sigma_y^2)$
Treatments	$t - 1$	σ_x^2	$\beta\sigma_x^2$	$\beta^2\sigma_x^2 + \sigma_y^2 + rT$
Error	$(r - 1)(t - 1)$	σ_x^2	$\beta\sigma_x^2$	$\beta^2\sigma_x^2 + \sigma_y^2$

The mean squares of y for treatments and error are equal in their expectations except for the T component in the former, and therefore the error mean square in an analysis of variance of this design leads to an unbiased estimate of the variance of the treatment means:

$$V(\bar{y}_i) = \frac{1}{r}(\beta^2\sigma_x^2 + \sigma_y^2). \tag{10}$$

Moreover, if an analysis of covariance is made according to the standard routine, using sums of squares and products for error and treatment-plus-error, the residual mean squares for treatment and error are again equal in expectation except for the T component. Hence the residual error mean square leads to an unbiased estimate of the average variance of a treatment mean:

$$V(\bar{y}_i') = \frac{1}{r}\sigma_y^2\left(1 + \frac{1}{rt - r - t - 1}\right). \tag{11}$$

5. COMPLETELY RANDOMIZED DESIGN (DESIGN I)

If the same N animals as in the randomized block design were allocated r to each treatment entirely at random, the variance between the blocks would be added proportionately into sums of squares and products for treatments and error, and the expected mean squares and products would be as in Table 2, where

$$h = t(r - 1)/(rt - 1). \tag{12}$$

Again an unbiased estimate of variance is obtainable from the error mean square for y, and

$$V(\bar{y}_i) = \frac{1}{r}[\beta^2\sigma_x^2 + \sigma_y^2 + h(B^2\Sigma_x^2 + \Sigma_y^2)]. \tag{13}$$

TABLE 2

EXPECTATIONS OF MEAN SQUARES AND PRODUCTS FOR COMPLETELY RANDOMIZED DESIGN (I)

Variation	D.f.	(x^2)	(xy)	(y^2)
Treatments	$t-1$	$\sigma_x^2 + h\Sigma_x^2$	$\beta\sigma_x^2 + hB\Sigma_x^2$	$\beta^2\sigma_x^2 + \sigma_y^2 + h(B^2\Sigma_x^2 + \Sigma_y^2) + rT$
Error	$t(r-1)$	$\sigma_x^2 + h\Sigma_x^2$	$\beta\sigma_x^2 + hB\Sigma_x^2$	$\beta^2\sigma_x^2 + \sigma_y^2 + h(B^2\Sigma_x^2 + \Sigma_y^2)$

The estimation of error again remains unbiased under an analysis of covariance. The expected value of the average variance of a treatment mean is a complicated expression, though its estimate is obtainable by ordinary analysis of covariance processes for any actual records.

$$V(\bar{y}_i') = \frac{1}{r}\left[\sigma_y^2 + h\Sigma_y^2 + \frac{h\sigma_x^2\Sigma_x^2(B-\beta)^2}{\sigma_x^2 + h\Sigma_x^2}\right]\left(1 + \frac{1}{rt - t - 2}\right). \quad (14)$$

Not only does it involve the y constituents of (13), in the manner that (11) is related to (10), but there is an additional item depending upon the difference between the two regression coefficients B and β, which will therefore often be zero.

If the animals for the experiment were selected independently at random from the infinite population, without regard to potential blocks, instead of being necessarily a set of N suitable for a randomized block design, Table 2 and equation (13) would be modified by writing $h = 1$. The change is slight when t and r are moderately large.

6. RANDOMIZED BLOCK DESIGN, BALANCED FOR TREATMENTS (DESIGN IV)

If N animals are randomly selected and allocated to animals with the dual constraint that they shall consist of r replicate blocks of t and that the t treatment means of x shall be equal, the analysis of covariance will be analogous to that in Table 1 but all constituents of the treatment line relating to x become zero. The total and block lines of the analysis are the same as if a true randomized block experiment were being conducted on the same animals, and consequently any reduction in the treatment line must be compensated by an increase in the error line, so keeping the sums of squares and products for treatment-plus-error unaltered. Table 3 shows the expected mean squares and products.

TABLE 3

EXPECTATIONS OF MEAN SQUARES AND PRODUCTS FOR RANDOMIZED BLOCK DESIGN BALANCED FOR TREATMENTS (IV)

Variation	D.f.	(x^2)	(xy)	(y^2)
Blocks	$r - 1$	$\sigma_x^2 + t\Sigma_x^2$	$\beta\sigma_x^2 + tB\Sigma_x^2$	$\beta^2\sigma_x^2 + \sigma_y^2 + t(B^2\Sigma_x^2 + \Sigma_y^2)$
Treatments	$t - 1$	0	0	$\sigma_y^2 + rT$
Error	$(r - 1)(t - 1)$	$\frac{r}{r-1}\sigma_x^2$	$\frac{r}{r-1}\beta\sigma_x^2$	$\frac{r}{r-1}\beta^2\sigma_x^2 + \sigma_y^2$

Doubtless many experimenters would consider that sufficient attention had been paid to x by balancing for this variate, and would therefore be content to make an analysis of variance on y alone. Comparison of the expected mean squares for treatment and error shows that their estimate of variance would be biased. The true variance of a treatment mean is

$$V(\bar{y}_i) = \sigma_y^2/r, \tag{15}$$

all dependence on x having been eliminated by the balancing, but an estimate uncritically formed from the error mean square would have as its expectation

$$\frac{\beta^2\sigma_x^2}{r-1} + \frac{\sigma_y^2}{r}.$$

The more marked the intra-block regression of y on x, the more successful is the balancing but the more serious is the consequent overestimation of error. The expectation of the biased estimate in fact exceeds the variance that is appropriate to design II.

The fault is very easily remedied. If the routine procedure for a covariance analysis is completed by the adjustment of y for x, the expected mean square for treatments is unaltered and that for error becomes simply σ_y^2. Thus unbiased estimation is achieved. Snedecor [1956] mentions that H. L. Lucas has reached a similar result in an unpublished paper. Because of the balance, the adjusted treatment means are the same as the unadjusted and

$$V(\bar{y}_i') = \sigma_y^2/r. \tag{16}$$

7. SIMPLE BALANCE FOR TREATMENTS (DESIGN III)

Table 4 is the modification of Table 2 required when the block classification of the N animals is ignored but treatment means of x are equalized; it bears to Table 2 the same relation that Table 3 does to Table 1.

Again the bias in the estimation of treatment variances that arises if y alone is analysed is apparent. The true variance of a treatment mean is

$$V(\bar{y}_i) = \frac{1}{r}[\sigma_y^2 + h\Sigma_y^2], \tag{17}$$

whereas division of the error mean square in Table 4 by r would give an estimate whose expectation also involves σ_x^2 and Σ_x^2 and exceeds that in (13). Adjustment for covariance on x leaves treatment means unaltered, because of the balancing, and therefore

$$V(\bar{y}_i') = \frac{1}{r}[\sigma_y^2 + h\Sigma_y^2]. \tag{18}$$

The residual mean square for error still leads to a biased estimate of this quantity, the variance of the means estimated from it having expectation

$$\frac{1}{r}\left[\sigma_y^2 + h\Sigma_y^2 + \frac{rt-1}{t(r-1)}\frac{h\sigma_x^2\Sigma_x^2(B-\beta)^2}{\sigma_x^2 + h\Sigma_x^2}\right]. \tag{19}$$

Hence variance estimation is biased unless either $\Sigma_x^2 = 0$ or $B = \beta$; the first is unlikely, since it would represent complete failure to stratify for x, but as already noted the second will be true in many experiments. Moreover, σ_x^2 will be relatively small in a successful stratification.

If the N animals were selected independently at random from the infinite population, Table 4 and (18) and (19) would be modified by writing $h = 1$.

8. COMPARISON OF DESIGNS

When stratification based on x has been well arranged, σ_x^2 , the variance within strata will be very small. In a randomized block design, little might be gained by a covariance adjustment for x unless the intra-block correlation of x and y were so large as to make $\beta^2\sigma_x^2$ comparable in magnitude with σ_y^2 ; this can be seen from (10) and (11). Indeed, in rare instances, if σ_x^2 were negligible the adjustment might do more harm than good, since the elimination of $\beta^2\sigma_x^2$ from (10) might be more than balanced by the extra factor in (11) that represents the average effect of the sampling errors of the adjustments. If by accident

TABLE 4

EXPECTATIONS OF MEAN SQUARES AND PRODUCTS FOR DESIGN WITH SIMPLE BALANCE FOR TREATMENTS (III)

Variation	D.f.	(x^2)	(xy)	(y^2)
Treatments	$t-1$	0	0	$\sigma_y^2 + h\Sigma_y^2 + rT$
Error	$t(r-1)$	$\frac{rt-1}{t(r-1)}(\sigma_x^2 + h\Sigma_x^2)$	$\frac{rt-1}{t(r-1)}(\beta\sigma_x^2 + hB\Sigma_x^2)$	$\frac{rt-1}{t(r-1)}(\beta^2\sigma_x^2 + hB^2\Sigma_x^2) + \sigma_y^2 + h\Sigma_y^2$

stratification has been less satisfactory, a covariance adjustment can still remove most of the ill effects, especially if r and t are large.

The inferiority of the completely randomized design scarcely requires comment, and is made clear by comparing (13) and (14) with (10) and (11). Without a covariance adjustment, the variance of treatment means is inflated by components of variance between the potential blocks. Even with adjustment, the variance contains a component depending upon such inter-block variance of y as is not associated with variations in x and a second component depending upon any difference between B and β; the final factor in (14) is smaller than that in (11), but this will rarely suffice to compensate for the other differences. Except perhaps for very small experiments, where the additional degrees of freedom for estimating error may be especially valuable, the completely randomized design will usually be less efficient than the randomized block. Under the conditions most favourable for complete randomization, namely $\Sigma_u^2 = 0$ and $B = \beta$, the true precision of this design relative to that in randomized blocks is

$$\frac{(N - t - 2)(N - r - t)}{(N - t - 1)(N - r - t - 1)},$$

a quantity only very slightly in excess of unity for any experiment with more than 20 animals. Making allowance for the further loss in precision when variances must be estimated from the experiment, by use of the formula (d.f. + 1)/(d.f. + 3) (Fisher [1951] §74) the efficiency becomes

$$\frac{(N - t - 2)(N - t)(N - r - t)(N - r - t + 3)}{(N - t - 1)(N - t + 2)(N - r - t - 1)(N - r - t + 1)},$$

which is slightly greater than the previous expression, but still too small to favour the completely randomized design. For 4 treatments tested on 5 animals each, it is only 1.065, and the 6.5% advantage can easily be balanced by a small value of Σ_u^2; for 2 treatments tested on 6 animals each, the value 1.38 gives more encouragement to complete randomization.

The chief interest of this paper lies with the two balanced designs. The argument of Sections 6 and 7 has shown that the variance estimates will be seriously biased unless a covariance analysis of y on x is used. For randomized blocks with treatment balance, the true variance of means is slightly less than that for ordinary randomized blocks; comparison of (16) with (11) shows the precision to be increased by a factor

$$\frac{N - r - t}{N - r - t - 1}.$$

In an experiment involving 20 animals, this may represent a gain of about 10%, and in an experiment on 40 animals about 4%. Similarly, for the completely randomized design the true variance of treatment means is smaller with treatment balance for x than without it, *if* B and β *are equal,* as is seen by comparing (18) and (14). The precision is increased by a factor

$$\frac{N - t - 1}{N - t - 2}$$

so that the gain is rather less than before. For either balanced design, covariance analysis is essential to unbiased estimation of variances.

One more comparison relevant to current practice is that between designs II and III. To some experimenters, the attraction of a simply balanced design is an avoidance of the need for randomized blocks. Whether or not the variance of unadjusted treatment means favours the balanced design depends upon the relative magnitudes of Σ_y^2 and $\beta^2\sigma_x^2$, as may be seen by comparison of (17) with (10); there is no basis for asserting the general superiority of III. If means are adjusted for covariance with x, comparison of (18) with (11) shows that unless Σ_y^2 is very small the advantage lies with randomized blocks. Even this argument presents the balanced design too favourably, for it ignores the impossibility of estimating (17) unbiasedly from the results of the experiment except when $B = \beta$; when this condition is satisfied, a covariance analysis leads to unbiased estimation, but if a covariance analysis is to be undertaken in any case, it might as well be done on a randomized block design where the results are likely to be still more precise.

Thus in terms of increase in precision of comparisons or decrease in variance of treatment means, the advantage for designs in which the treatments are deliberately equalized in respect of mean values of the initial measurement is small, except for experiments involving less than 20 animals in all. Moreover, the analysis here has assumed that perfect balance is attained; any inequality in the means of x for the treatment groups will reduce the precision to an intermediate value. Against any advantage must be set either the technical difficulties of arranging the balance objectively or the theoretical objections to the balance usually achieved on the grounds of its subjective character. In the writer's opinion, balance of this kind is very seldom worth the price that is paid for it, and he believes that most experimenters who are attracted by its apparent virtues would be wiser to abandon it. In special circumstances, the decision between balanced designs and orthodox randomized designs should perhaps rest with the experimenter,

but he needs to be aware of the main conclusion from statistical theory that he cannot greatly increase his precision by adopting balance. The statistician may perhaps be forgiven for wondering why there should be anxiety to avoid a design that is as simple as randomized blocks, especially when the considerations advanced above show that, even under the most favourable circumstances, the advantage for a balanced design must be small and often it will be substantially inferior.

A possibility of attaining a degree of balance in respect of x closer than will usually occur with randomized blocks, yet without the dangers of bias inseparable from designs III and IV, may finally be noted. If r is an integral multiple of t, the animals might first be allocated to r blocks as for randomized blocks, then arranged in descending order of x within each block, after which the treatments could be assigned in respect of blocks and order within blocks according to r/t randomly selected Latin squares. The standard Latin square analysis would then eliminate from the comparisons of treatment means for y any additional component of variation associated with order within blocks. Covariance analysis might still lead to further gain in precision. If r/t is not an integer, Youden squares and other incomplete Latin square schemes might be employed, at some considerable cost in respect of computation-time. In practice, probably this Latin square constraint will seldom be very advantageous, but it has the merit of increasing precision without a covariance analysis in the same circumstances that would make design IV more precise than simple randomized blocks. The loss of information arising from the fewer error degrees of freedom in the Latin square will scarcely be important except in very small experiments (e.g. a single 4×4 Latin square) in which the additional constraint is completely without effect.

9. SUMMARY

Experimenters sometimes impose on the animals or other experimental units allocated to different treatments the condition that treatment means in respect of some preliminary measurement, believed to be correlated with the measurement that is to be used in assessment of the results, shall be approximately equal. Such balance may be incorporated into completely randomized or randomized block designs. The expected variances of treatment means for these designs and for the corresponding randomized designs without balance have been studied and compared. Three major criticisms of the balanced designs emerge:

(i) It is difficult to define an objective plan of allocation of units to treatments that will fulfill the requirements of balance yet allow the particular allocation adopted to be regarded as a

random selection from a number of possibilities, and so give some hope that estimates of means and variances unbiased by subjective influences can be obtained.

(ii) Even though the dangers of (i) be disregarded, uncritical analysis of balanced design will seriously overestimate the variances of treatment means. By use of a covariance analysis with the initial measurement, all or most of this bias can be removed, but this *essential* step in the analysis of a balanced design destroys much of its advantage in respect of simplicity and and speed of calculation.

(iii) When (i) has been disregarded and the covariance calculation demanded by (ii) completed, there still remains the question as to whether any gain in precision has been achieved. The answer is that, except possibility for very small experiments, the gain is at best of the order of a few percent and is almost always too small to be worth setting against the risk of bias under (i) and the loss in precision that can occur if conditions are not optimal. More particularly, a randomized block design in which the blocks are based upon the known values of the initial measurement will almost certainly be more precise than a design in which blocks are abandoned in favour of balance.

Thus the disadvantages of balance of this kind seem far to outweigh the advantages, and such designs should not be adopted unless special circumstances dictate their advisability.

The possibility of using Latin square and allied designs as alternatives to the type of balance discussed has been briefly mentioned. When there is anything to be gained by balance, these designs avoid the theoretical objections and will usually have almost equal merits.

REFERENCES

Finney, D. J. [1946] Standard errors of yields adjusted for regression on an independent measurement. *Biometrics, 2:* 53–55.

Fisher, R. A. [1951] *The Design of Experiments* (6th edition). Oliver and Boyd, Ltd., Edinburgh.

Kempthorne, O. and Wilk, M. B. [1955] Fixed, mixed, and random models. *Journal of the American Statistical Association, 50:* 1144–1167.

Snedecor, G. W. [1956] *Statistical Methods* (5th edition). Iowa State College Press, Ames, Iowa.

THE ANALYSIS OF COVARIANCE AS A MISSING PLOT TECHNIQUE

Irma Coons
Westinghouse Electric Corporation, East Pittsburgh, Pa., U.S.A.

INTRODUCTION

The occurrence of missing data in a statistically designed experiment requires some modification of the routine technique of analysis. For certain simple designs, such as randomized blocks and Latin squares, formulae are readily available by which a single missing value may be estimated, and the single missing value procedure may be used iteratively to estimate several missing values. The analysis of the augmented data may then be completed as usual after certain corrections for bias are made. One of these—a decrease in the degrees of freedom associated with total and with error sources of variation—is easily made. The other correction is necessitated by the fact that the treatment sum of squares, as calculated from the augmented data, is incorrect. In the case of a single missing value the treatment sum of squares is too large. It can be shown that this holds with any number of missing values. The general procedure for getting a correct treatment sum of squares can of course be applied.

The work required to estimate the missing values and the resulting bias may become very tedious. Also, situations may arise for which no general formula is available. For instance, a fractionally replicated design does not always use the same effects to obtain an estimate of residual error. Since an estimate of a missing value is a function of the error term, a formula would have to be recalculated for each problem. Also, missing observations in split plot designs may cause problems. Some general method of handling missing observations in any situation is a needed tool.

The purpose of this paper is to illustrate the full details of a method which can be used when one or more missing observations exist in an experiment of any statistical design. The advantages of this method are its generality of application and the ease with which exact tests of significance may be obtained. Here the word "exact" is used to mean exact when errors are normally and independently distributed. The technique employs the computational procedures of a covariance

analysis using a dummy X covariate. Originally given in a paper by Bartlett [1937] and also described by Anderson [1946], the method does not seem to have been exploited to the full possible advantage and merits further description.

GENERAL USE OF COVARIANCE TO DEAL WITH MISSING DATA

In this section some properties will be listed which give the justification for the computational procedures to be described. Readers who are interested only in the computational application of the missing plot technique may omit this section. Some of the following statements are easily proved; others are more difficult except with matrix notation. The properties shown here serve to clarify some obscurities in missing plot techniques discussed by Norton [1955], Nelder [1954], and Smith [1957]. Property 1 is due to Fisher, property 2 has been implicitly assumed by many authors, property 3 is due to Bartlett [1937], properties 4, 5, 6 have been obtained by Kempthorne (unpublished), but may be known to a number of workers.

1. If an analysis of variance is made with symbols $\beta_1 , \beta_2 , \cdots \beta_q$ in the place of missing observations, then the best linear unbiased estimates of the missing observations are the quantities $\hat{\beta}_1 , \hat{\beta}_2 , \cdots \hat{\beta}_q$ which minimize the error sum of squares.

2. Given that, with full data $(y_1 , y_2 , \cdots y_n)$, the best linear unbiased estimate of some linear function of the parameters is $v_1y_1 + v_2y_2 + \cdots + v_ny_n$, then the best estimate of that function with missing data is obtained by replacing the missing y's with the missing value estimates.

3. Let the data be observed data where obtained and zero where missing. Introduce a concomitant variable $X_m (m = 1 \cdots q)$ corresponding to the mth missing observation; let X_m take the value $-v$ for the mth missing observation and zero for all others, missing or not. If the error partial regression coefficients obtained from an analysis of covariance are denoted by $\hat{\beta}_1 , \hat{\beta}_2 , \cdots \hat{\beta}_q$, then $v\hat{\beta}_1 , v\hat{\beta}_2 , \cdots v\hat{\beta}_q$ are the best linear unbiased estimates of the missing observations.

4. Estimates of functions of data with missing observations, and variances and covariances of these estimates may be obtained by the routine application of formulae for adjusted means in the analysis of covariance; i.e. by regarding the zero yields supplied in the analysis of covariance procedure as having variances of σ^2. The above statement applies to functions of the augmented data; the variance of a missing observation *per se* is given by statement 5 following.

5. Denote the error sum of squares of X_i by E_{ii} and the error sum of products of X_i and X_j by E_{ij} . Then the variance of the ith missing

value estimate is $(v^2u_{ii} - 1)\sigma^2$ and the covariance of the ith and jth missing value estimates is $v^2u_{ij}\sigma^2$, where

$$\begin{bmatrix} E_{11} & E_{12} \cdots E_{1q} \\ E_{21} & \\ \vdots & \\ E_{q1} & E_{qq} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \cdots u_{1q} \\ u_{21} & \\ \vdots & \\ u_{q1} & u_{qq} \end{bmatrix} = \begin{bmatrix} 1 & & \\ & & 0 \\ & 0 & \\ & & 1 \end{bmatrix}$$

6. The sum of squares for treatments obtained by analyzing the data augmented by the missing value estimates is always greater than or equal to the exact sum of squares for treatments.

APPLICATION OF COVARIANCE TECHNIQUE TO ONE MISSING OBSERVATION

Let n equal the number of observations in the experiment including the missing one. Consider the original data as the Y variable of the covariance analysis and insert the value of zero in the cell which has the missing observation. Set up an X variable which takes the value of $-n$ in the cell corresponding to the substituted zero value and the value of zero elsewhere. The choice of $-n$ for the concomitant variable leads to the simple consequence that the sum of squares for any concomitant variable is equal to $n \times$ (degrees of freedom), as described below.

The usual computational procedures of the covariance analysis automatically provide unbiased tests of significance. However estimates of functions of the Y data, such as treatment means, must be adjusted to the value of zero for the concomitant variable rather than to the observed average value of X as is usually done. For example, an adjusted treatment mean is not estimated by $\bar{Y}_{.j} - \hat{\beta}(\bar{X}_{.j} - \bar{X}_{..})$, the covariance formula generally used, but rather by the formula:

$$\text{adj} \quad \bar{Y}_{.j} = \bar{Y}_{.j} - \hat{\beta}\bar{X}_{.j} \tag{1}$$

where

$$\hat{\beta} = \frac{E_{xy}}{E_{xx}},$$

and E_{xy} and E_{xx} are the error sum of products and sum of squares respectively, in the analysis of covariance. With a hierarchical classification, the adjustment is to the average of the next higher classification until the highest classification is reached, and the adjustment is then to zero, as shown above.

The variance of (1) is given by:

$$V(\text{adj } \bar{Y}_{.i}) = V(\bar{Y}_{.i}) + (\bar{X}_{.i})^2 V(\hat{\beta}) \tag{2}$$

where $V(\hat{\beta})$ is taken to be σ^2/E_{xx}, and σ^2 is the true variance of any observation. This is estimated by replacing σ^2 by s^2, where s^2 is the residual error mean square resulting from routine application of the analysis of covariance. That is,

$$s^2 = \frac{1}{df}\left(E_{yy} - \frac{E_{xy}^2}{E_{xx}}\right),$$

where E_{yy} is the covariance error sum of squares of the y's, and the degrees of freedom are equal to the error degrees of freedom with full data, less one.

Any comparison among the adjusted treatment means, such as:

$$\text{adj } T_m - \text{adj } T_n = (\bar{Y}_{.m} - \bar{Y}_{.n}) - \hat{\beta}(\bar{X}_{.m} - \bar{X}_{.n})$$

may in general be designated by the notation

$$C = C_y - \hat{\beta}C_x \,. \tag{3}$$

The variance of such a comparison is:

$$V(C_y) + C_x^2 V(\hat{\beta}), \tag{4}$$

where $V(C_y)$ is calculated assuming no missing observations and $V(\hat{\beta})$ is taken to be σ^2/E_{xx}.

As noted in statement 3 above,

$$\text{Missing observation} = n\hat{\beta} = n\frac{E_{xy}}{E_{xx}}.$$

From statement 5 above, the variance of $n\hat{\beta}$ is

$$\sigma^2\left(\frac{n^2}{E_{xx}} - 1\right).$$

However the complete analysis of the data may be performed without estimating the missing observation as such.

The approximate test of significance for any source of variation, say T, is obtained by computing the biased sum of squares of T, as

$$T_{yy} - 2\hat{\beta}T_{xy} + \hat{\beta}^2 T_{xx} \,,$$

where T_{yy}, T_{xy}, and T_{xx} are respectively the sum of squares for y, of products for x and y, and of squares for x, for the particular source T.

Of course, when there is a missing value formula available for the particular case under consideration, this approximate sum of squares

can also be obtained by estimating the missing value by this formula and then analyzing the augmented data. However, the virtue of the covariance technique is that only minor supplementary computations make the exact test readily obtainable as the ordinary unbiased analysis of covariance test. This is exemplified in the examples given.

The additional computation required by the use of an X covariate is relatively slight, due to the simple nature of the X data. In most situations, the corrected sum of squares of X attributable to any given source of variation ($\sum x_i^2$) will simply be

$$n \times \text{(degrees of freedom for the given source of variation)}.$$

If, of course, one were analyzing the data by a high speed computing device, there would be no point in working out the partition of the sum of squares of x or products of x and y algebraically, since the cost of doing this on the device would generally be negligible. From this point of view, the procedure seems preferable as a universal missing plot procedure to that proposed by Hartley [1956].

It is also possibly worth mentioning that this covariance device is one extremely effective method for obtaining missing plot formulae. The reader may verify that the derivation of missing plot formulae for randomized blocks and Latin squares, for example, is entirely trivial by the covariance procedure. For incomplete block designs, the covariance procedure is relatively simple to apply, whereas a missing value formula could be complex.

EXAMPLES OF COVARIANCE APPLIED TO ONE MISSING OBSERVATION

Fractionally Replicated Design of a $2^5 \times 4$ Factorial

The data used to illustrate this type of design are a part of a larger experiment run at Westinghouse Electric Corporation. The experiment was designed to investigate the effects of various processing variables upon the corrosion properties of ZIRCALOY. The measurement used is weight gain per unit surface area, and the factors considered in Table 1 are listed below. These treatments were run in factorial combination but only a 1/4 replicate was used.

Position in Autoclave (A)	1. 2 (Lowest level tested)
	2. 4
	3. 6
	4. 8
Control of vacuum in Annealing (B)	1. Minimum leak rate
	2. Constant leak rate

Cleaning treatment, prior to pickling (C)	1. An alkaline cleaner
	2. A vapor degreaser
Acidity of the first rinse, after pickling (D)	1. pH 6–8
	2. pH 2–3
Method of the second rinse, after pickling (E)	1. Sprayer
	2. Hose
Acidity of the third rinse (F)	1. pH 6–8
	2. pH 4–5

The missing observation is the treatment combination $a_3b_2c_2d_2e_1f_1$, as shown in Table 1. The X values are not shown for lack of space, but they would simply consist of $-32 = -n$ in the cell X_{322211} , and of zeros elsewhere.

The covariance analysis of the data is given in Table 2, with all the computations relative to a given treatment listed on a single line. All factors are considered fixed, not random.

The computations in this experiment may appear long, but are quite simple. The X analysis of column (3) may be written down at sight since

$$\sum x_j^2 = n \times (\text{degrees of freedom}).$$

For the factors and interactions of column (2) with a single degree of freedom, the sum of products of x and y is simply

$$Y_1 - Y_2$$

where Y_1 = total of Y observations for that level of the effect or interaction which does *not* contain the missing observation and Y_2 = total for that level which does contain the missing observation. Also for these single degree of freedom effects and interactions, the sum of squares for y is simply

$$(Y_1 - Y_2)^2/n.$$

Column (4) of Table 2 gives those approximate mean squares which would be obtained by an analysis of the data augmented by the estimated missing observation. This is equivalent to an analysis of

$$Y - \hat{\beta}X$$

so the approximate sum of squares may be calculated as

$$\sum y^2 + \hat{\beta}^2 \sum x^2 - 2\hat{\beta} \sum xy.$$

From statement 6, these approximate mean squares are known to be too large. Therefore any of these which are not significant may be eliminated from further consideration, and the calculation required in the analysis may thus be shortened.

TABLE 1

DATA FOR ONE-FOURTH REPLICATION OF A $2^5 \times 4$ COMPLETELY RANDOMIZED FACTORIAL ZIRCALOY PROCESSING INVESTIGATION

Control of vacuum in annealing (*B*)			Minimum leak Rate				Constant leak rate			
Cleaning treatment (before pickling) (*C*)			Alkaline		Vapor degreaser		Alkaline		Vapor degreaser	
Acidity of 1st rinse (*D*)	Method of 2nd rinse (*E*)	Acidity of 3rd rinse (*F*)	Observation	Position in autoclave (*A*)	Observation	Position in autoclave (*A*)	Observation	Position in autoclave (*A*)	Observation	Position in autoclave (*A*)
*p*H 6–8	Sprayer	*p*H 6–8	28	4	33	1	30	1	30	4
		*p*H 4–5	34	2	32	3	28	3	35	2
	Hose	*p*H 6–8	31	2	32	3	30	3	33	2
		*p*H 4–5	31	4	34	1	33	1	32	4
*p*H 2–3	Sprayer	*p*H 6–8	30	3	37	2	45	2	00	3
		*p*H 4–5	37	1	39	4	46	4	44	1
	Hose	*p*H 6–8	33	1	36	4	40	4	38	1
		*p*H 4–5	35	3	37	2	47	2	41	3

TABLE 2

COVARIANCE ANALYSIS OF A ONE-FOURTH REPLICATION OF A COMPLETELY RANDOMIZED DESIGN WITH ONE MISSING OBSERVATION

Source of variation	D.f.	(1) $\sum y^2$	(2) $\sum xy$	(3) $\sum x^2$	(4) Approx. mean sq.	(5) Treatment + Error $\sum y^2$	(6) $\sum xy$	(7) $\sum x^2$	(8) $\sum y^2 - \frac{(\sum xy)^2}{\sum x^2}$	D.f.	(9) Adjusted treatment mean sq.
Type of vacuum (B)	1	5.28	−13	32	82.56*	627.71	486	448	100.49	1	76.62*
Cleaning treatment (C)	1	19.53	25	32	5.61						
$B \times C$	1	140.28	67	32	25.56*	762.71	566	448	47.63	1	23.76*
Acidity of 1st rinse (D)	1	195.03	−79	32	430.71*	817.46	420	448	423.71	1	399.84*
$B \times D$	1	13.78	−21	32	110.26*	636.21	478	448	126.20	1	102.33*
$C \times D$	1	101.53	57	32	10.81*	723.96	556	448	33.92	1	10.05*
Method of 2nd rinse (E)	1	38.28	35	32	0.36						
$B \times E$	1	42.78	37	32	0.06						
$C \times E$	1	30.03	31	32	1.71						
$D \times E$	1	16.53	23	32	7.41						
Acidity of 3rd rinse (F)	1	195.03	79	32	51.51*	817.46	578	448	71.74	1	47.87*
$B \times F$	1	52.53	41	32	0.21						
$C \times F$	1	30.03	31	32	1.71						
$D \times F$	1	94.53	55	32	8.61						
$E \times F$	1	63.28	45	32	1.36						
Position in autoclave (A)	3	357.84	179	96	22.16*	980.27	678	512	82.45	3	19.53*
Residual	13	622.43	499	416					23.87	12	$2.00 = s_E^2$

Estimate of $\hat{\beta} = \frac{Exy}{Exx} = 1.20$

Estimate of missing observation $= n\hat{\beta} = 32\ (1.20) = 38$

Since all factors are considered fixed, the residual variation is used as the error term throughout. Therefore the (Treatment + Error) columns are obtained by simply adding the same residual term to each significant treatment effect in the column; similarly the constant adjusted residual, 23.87, is subtracted from each line of column (8) to obtain adjusted treatment sum of squares.

Thus the completed analysis is obtained and exact tests of significance have been made as shown. One table of means, that of the interaction between the effects of Type of Vacuum × Acidity of First Rinse ($B \times D$) has been presented (Table 3) as an example of the procedure which must be followed to obtain the appropriate means and their variances.

TABLE 3

TABLE OF MEANS

TYPE OF VACUUM × ACIDITY OF FIRST RINSE ($B \times D$) INTERACTION

Type of vacuum (D)	Acidity of the first rinse (B)	
	pH 6–8	*pH* 2–3
Minimum leak rate	32	31
Constant leak rate	36	42

Missing observation occurred under the Constant Leak Rate Vacuum at *pH* 2–3.

$$\text{Adj } \bar{Y}_{.2\,2..} = \bar{Y}_{.2.2..} - \hat{\beta}\bar{X}_{.2.2..}$$
$$= 37.6 - (1.2)(-4) = 42$$
$$V(\text{Adj } Y_{.2.2..}) = s_E^2 [1/8 + (-4)^2/416] = .33$$
$$V(Y_{i\,k\,.},\ i,\ k \neq 2) = s_E^2 (1/8) = .25$$

Split-plot Design With One Missing Observation

Another phase of the ZIRCALOY experiment was concerned with the effects of production sources of variation upon corrosion of ZIRCALOY. The factors here considered are:

Sponge producer (S)	1. Producer 1
	2. Producer 2
Time (days) (T)	1. 14
	2. 28
	3. 42
	4. 84
Position in the autoclave (P)	1. Low level
	2. High level

Duplicate readings were taken on each treatment combination. The factors were run in a split-plot design with positions as the subplot. Data given in Table 4 are a function of the logarithms of the weight gain per unit area. The subplot taken as missing is associated with treatment combination $t_1 s_2 p_2$. Again the X data are not shown but may be easily visualized.

The covariance analysis is given in Table 5. The new feature of this design is the presence of more than one error term. Actually, two estimates of the missing observation are available—one ($n\hat{\beta}_E$) applicable to the individual subplot and one ($n\hat{\beta}_W$) to the whole plot which contains the missing subplot. This purely technical point may be worrisome to the non-statistician but a fully adjusted set of means should remove the difficulty. The procedure by which the mean of every whole plot and split-plot treatment combination may be written down is given by Kempthorne [1952, p. 388]. It is to be noted, however, that the adjustment is to a zero value of the concomitant. One finds that existent observations occurring in a whole plot with missing sub-plots are adjusted in such a way as to retain all comparisons given by the existent data.

The analysis and tests of treatment effects proceed as usual for a split-plot covariance analysis. A linear corrosion rate per day, a main plot effect, is given by the contrast

$$\frac{1}{196}\left[-2\bar{Y}_1 - \bar{Y}_2 + 3\bar{Y}_4 - \hat{\beta}_W(-2\bar{X}_1 - \bar{X}_2 + 3\bar{X}_4)\right]$$

$$= \frac{1}{196}\left[953 - 16.7(8)\right] = 4.18$$

(From Equation 3)

and the variance of this contrast is

$$\frac{1}{196^2}\left[V(C_v) + V(\hat{\beta}_W)C_x^2\right] = \frac{1}{196^2}\left[\frac{14s_W^2}{8} + \frac{s_W^2}{W_{xx}}8^2\right] = .27$$

(From Equation 4)

The adjusted difference between high and low positions in the autoclave, a subplot effect, is estimated by

$$\bar{Y}_1 - \bar{Y}_2 - \hat{\beta}_E(\bar{X}_1 - \bar{X}_2) = [74 - 14.0(2)] = 46 \quad \text{(From Equation 3)}$$

with a variance of

$$V(C_v) + V(\hat{\beta}_E)C_x^2 = \left[\frac{2s_E^2}{16} + \frac{s_E^2}{E_{xx}}2^2\right] = 28 \quad \text{(From Equation 4)}$$

TABLE 4

DATA FROM SPLIT-PLOT DESIGN FOR ZIRCALOY PRODUCTION INVESTIGATION

Sponge producer (S)	Producer 1				Producer 2				Total for positions
Position in autoclave (P)	Time (days) (T) 14	28	42	84	Time (days) (T) 14	28	42	84	
Low	518	591	716	839	462	568	681	820	
	613	681	740	799	505	623	748	875	10779
High	447	518	634	799	000	518	623	799	
	568	663	690	763	491	568	681	826	9588
Total for time	2146	2453	2780	3200	1458	2277	2733	3320	

TABLE 5

COVARIANCE ANALYSIS OF SPLIT-PLOT DESIGN WITH ONE MISSING OBSERVATION

Source of variation	D.f.	$\sum y^2$	$\sum xy$	$\sum x^2$	Treatments plus error $\sum y^2$	$\sum xy$	$\sum x^2$	$\sum y^2 - \frac{(\sum xy)^2}{\sum x^2}$	Adjusted treatment sum of squares	D.f.	Adjusted treatment mean square	F-Ratio
Time (T)	3	570202	5951	32(3)	678348	10223	32(11)	381446	344588	3	114863	**
Sponge producer (S)	1	19553	791	32	126700	5063	32(9)	37692	835	1	835	–
$S \times T$	3	45564	1961	32(3)	153711	6233	32(11)	43340	6482	3	2161	–
Error (W)	8	108147	4272	32(8)				36858		7	$5265 = s_W^2$	
Position (P)	1	44328	1191	32	95912	4775	32(9)	16742	15335	1	15335	**
$P \times T$	3	15203	1177	32(3)	66787	4761	32(11)	2392	984	3	328	–
$P \times S$	1	4073	361	32	55657	3945	32(9)	1618	210	1	210	–
$P \times T \times S$	3	12163	1079	32(3)	63747	4663	32(11)	1975	567	3	189	–
Error (E)	8	51584	3584	32(8)				1408		7	$201 = s_E^2$	

$n\hat{\beta}_W = 534$
$n\hat{\beta}_E = 448$

APPLICATION OF COVARIANCE TECHNIQUE TO MORE THAN ONE MISSING OBSERVATION

With more than one missing observation a multiple covariance analysis is required. Again let n equal the number of Y observations in the experiment including the missing ones. Assign a value of zero to those Y observations which are missing. Set up a concomitant variable X_m for each missing observation. Each of these X_m will consist of zero in all cells except in that cell corresponding to the missing observation with which the given X_m is associated; in that one cell it will have a value of $-n$.

With q missing observations, a multiple covariance analysis must be performed on Y and the q covariates X_m. An outline of a general multiple covariance analysis is given in Table 6 for a randomized block experiment with two missing observations ($q = 2$). The computations required to obtain the $\sum x_m x_n$ and $\sum x_m y$ are simple, since each X_m is associated with a single missing value and therefore has only one non-zero cell. In computing $\sum x_m x_n$, two situations may be encountered.

1. The two missing values associated with X_m and X_n occur in the same level of the given source of variation. The results then are exactly the same as those obtained for $\sum x^2$; i.e.

$\sum x_m x_n = n \times$ (degrees of freedom for the given source of variation) in most cases.

2. The two missing values occur in different levels of the given source of variation. Then, for most cases,

$$\sum x_m x_n = -nr$$

where r is dependent upon the hierarchical classification which is used. When no hierarchical classification is present, $r = 1$.

When the given source of variation is an interaction effect, then the corresponding main effects and lower order interactions must be subtracted from the above $\sum x_m x_n = -nr$.

A complete analysis of covariance will provide unbiased tests for treatment effects. Each missing observation Y_m may be estimated by $n\hat{\beta}_{m_E} = n \times$ (error estimate of the $\hat{\beta}$ associated with that missing observation).

If q observations are missing, their estimation requires the solution of a set of simultaneous equations of size $q \times q$ in order to obtain estimates of the $\hat{\beta}_{m_E}$, as shown in Table 6. Each unbiased test of adjusted treatment effects requires the solution of an additional set of simultaneous equations of size $q \times q$. In general, the solution of these equations will be quite easy, even by an iterative technique, because usually the off-diagonal terms of the matrix of coefficients will be small.

TABLE 6

ANALYSIS OF MULTIPLE COVARIANCE IN A RANDOMIZED BLOCK DESIGN TWO COVARIATES

Source of variation	D.f.	$\sum x_1^2$	$\sum x_1x_2$	$\sum x_2^2$	$\sum x_1y$	$\sum x_2y$	$\sum y^2$	$\sum y^2 - \hat{\beta}_1\sum x_1y - \hat{\beta}_2\sum x_2y$	D.f.	Mean square
Blocks	$(b-1)$	$B_{x_1x_1}$	$B_{x_1x_2}$	$B_{x_2x_2}$	B_{x_1y}	B_{x_2y}	B_{yy}			
Treatments	$(t-1)$	$T_{x_1x_1}$	$T_{x_1x_2}$	$T_{x_2x_2}$	T_{x_1y}	T_{x_2y}	T_{yy}			
Error	$(b-1)(t-1)$	$E_{x_1x_1}$	$E_{x_1x_2}$	$E_{x_2x_2}$	E_{x_1y}	E_{x_2y}	E_{yy}	S_E	$(b-1)(t-1)-2$	$S_E/df = s_E^2$
Treatment + error	$b(t-1)$	$T_{x_1x_1}+E_{x_1x_1}$	$T_{x_1x_2}+E_{x_1x_2}$	$T_{x_2x_2}+E_{x_2x_2}$	$T_{x_1y}+E_{x_1y}$	$T_{x_2y}+E_{x_2y}$	$T_{yy}+E_{yy}$	S_{T+E}	$b(t-1)-2$	
Adjusted treatments								$S_{T+E}-S_E$	$t-1$	$(S_{T+E}-S_E)/df$

Estimates for β_{1_E} and β_{2_E}, to be used in obtaining S_E, are given by solving the system of equations

$$E_{x_1x_1}\beta_{1_E} + E_{x_1x_2}\beta_{2_E} = E_{x_1y}$$
$$E_{x_2x_1}\beta_{1_E} + E_{x_2x_2}\beta_{2_E} = E_{x_2y}$$

Estimates for $\beta_{1_{T+E}}$ and $\beta_{2_{T+E}}$, to be used in obtaining S_{T+E}, are given by solving the system of equations:

$$(T_{x_1x_1}+E_{x_1x_1})\beta_{1_{T+E}} + (T_{x_1x_2}+E_{x_1x_2})\beta_{2_{T+E}} = T_{x_1y}+E_{x_1y}$$
$$(T_{x_2x_1}+E_{x_2x_1})\beta_{1_{T+E}} + (T_{x_2x_2}+E_{x_2x_2})\beta_{2_{T+E}} = T_{x_2y}+E_{x_2y}$$

Test for Treatment Effects $F_{(b-1)(t-1)-2}^{(t-1)} = (S_{T+E}-S_E)/(t-1)s_E^2$

Even so, the amount of computations required may be reduced by performing an approximate analysis of variance on the Y data augmented by the missing value estimates $n\hat{\delta}_{mx}$, as was done in Table 2 with a single missing observation. Since these approximate treatment sums of squares are known to be biased upward (see statement 6), any treatment effects which are not significant in this approximate analysis will be non-significant in an exact analysis; exact tests may be made on significant treatment effects if desired. With the above procedure the missing values are estimated exactly without tedious iteration, and the approximate analysis is available from the solution of a single $q \times q$ set of simultaneous equations.

A virtue of the covariance technique whenever split-plots occur is that the various cases of missing split-plots in the same or different whole-plots are taken care of automatically. If, for instance, two missing split-plots occur in the same whole plot, the two corresponding concomitants will have perfect correlation for the whole-plot analysis. When they are thus identical in sets, only one from each set should be taken into account in finding the whole-plot error regression. In such a situation only the sum of the regression coefficients ($\hat{\beta}_m$) associated with these concomitants can be estimated. However, because of the perfect correspondence in whole plot totals, this sum is all that is needed to obtain adjusted Y values. For instance, with two missing split-plots in the (ij) whole-plot, the adjustment process in the simple split-plot situation for split-plot (ijk) is:

$$y_{ijk} - \hat{\beta}_{1w}\bar{x}_{1ij.} - \hat{\beta}_{2w}\bar{x}_{2ij.} - \hat{\beta}_{1e}(\bar{x}_{1ijk} - \bar{x}_{1ij.}) - \hat{\beta}_{2e}(\bar{x}_{2ijk} - \bar{x}_{2ij.}).$$

But $\bar{x}_{1ij.}$ equals $\bar{x}_{2ij.}$ so that we need only $(\hat{\beta}_{1w} + \hat{\beta}_{2w})$ and the routine procedure provides this quantity. The above adjustment formula shows that all split-plot observations in the whole-plot containing missing values are subjected to an adjustment; however, this adjustment is such that the split-plot comparisons as given by the actual data are retained.

EXAMPLES OF COVARIANCE APPLIED TO MORE THAN ONE MISSING OBSERVATION

Split-Plot Design With Two Missing Subplots From Different Whole Plots

These data are taken from a larger experiment investigating the effects of various production factors upon changes in energy loss as raw steel is made into transformers. The factors shown in Table 7 are:

Coating used on the sheet of steel (A)
1. Glass
2. Carlite

Location from which strip was taken (B)	1. Edge of sheet 2. Center of sheet
Tension under which strips were wound into rings (C)	1. Low 2. High

The factor A, replicated four times in each level, forms the main plot. The split plots are the four treatment combinations of factors B and C. The missing values are treatment combinations $a_1b_2c_1$ in replicate 4, and $a_2b_1c_1$ in replicate 3. For compactness the tables of X_1 and X_2 are condensed into a single table; actually each table should consist of a single $-32 = -n$ in the appropriate cell and zero elsewhere.

After obtaining estimates of the missing values, an approximate analysis was performed upon the data augmented by main plot estimates

TABLE 7

DATA FOR THE SPLIT-PLOT DESIGN WITH TWO MISSING OBSERVATIONS

Coating used on the steel (A)		Glass				Carlite			
		R_1	R_2	R_3	R_4	R_1	R_2	R_3	R_4
Location of origon of strips (B)	Tension (C)	*Y Data*							
Edge	Low	75	69	66	56	42	75	00	62
	High	70	51	68	50	59	66	106	102
Center	Low	122	91	102	00	39	50	74	34
	High	114	125	84	72	82	46	77	56
		X Data							
Edge	Low	00	00	00	00	00	00	$-32 = X_2$	00
	High	00	00	00	00	00	00	00	00
Center	Low	00	00	00	$-32 = X_1$	00	00	00	00
	High	00	00	00	00	00	00	00	00

when computing main-plot effects A and Error (W), and by subplot estimates elsewhere (Table 8). In the split-plot design there are two approximate tests for whole-plot effects—one obtained solely from the main-plot analysis with data augmented by the main-plot estimates $n\hat{\beta}_{1W}$, $n\hat{\beta}_{2W}$, $\cdots$, and the second based on the analysis of the data augmented by the split-plot estimates $n\hat{\beta}_{1E}$, $n\hat{\beta}_{2E}$, $\cdots$. In order to obtain the correct subplot error term in the approximate analysis, subplot missing value estimates must be used in obtaining the main plot as well as subplot effects of the analysis. In Table 8, for instance, if the AB interaction effect were to be computed by subtracting A and B main effects from the raw corrected AB sum of squares, then the estimate of A which must be used is not 5832 as shown for the main plot analysis, but rather $(1215 + 84 - 970 - 86)^2/32 = 1847$, which is the A sum of squares using subplot missing value estimates. Therefore, the easiest method of obtaining the subplot error terms for the approximate analysis seems to be by the usual covariance formula

$$E_{yy} - \sum_{m} \hat{\beta}_{mE}(E_{xmy}),$$

since the error terms in the covariance analysis and approximate analysis must be the same and since the $\hat{\beta}_{mE}$ have already been calculated.

In the approximate analysis of Table 8, the effects of A and AB are significant. Exact tests for these effects are then shown. In order to obtain the exact sum of squares for A, for instance, the following adjustment must be applied to Effect A + Error (W):

$$\sum_{m=1}^{2} \hat{\beta}_{mA+W}(A + W)_{xmy}$$

where the $\hat{\beta}_{mA+W}$ must satisfy the equations (see Table 6):

$$7(32)\hat{\beta}_{1A+W} - 32\hat{\beta}_{2A+W} = 761$$

$$-32\hat{\beta}_{1A+W} + (7)32\hat{\beta}_{2A+W} = 129$$

Therefore

$$\hat{\beta}_{1A+W} = 3.55$$

$$\hat{\beta}_{2A+W} = 1.08.$$

Hence the adjusted [A + Error (W)] sum of squares is

$$\sum y^2 - \sum_{m} \hat{\beta}_{mA+W}(A + W)_{xmy}$$

$$= 7845 - 3.55(761) - 1.08(129) = 5004.$$

TABLE 8

ANALYSIS OF SPLIT-PLOT DESIGN WITH TWO MISSING OBSERVATIONS

Source of variation	D.f.	$\sum x_m^2$	$\sum x_1x_2$	$\sum x_1y$	$\sum x_2y$	$\sum y^2$	Approx. $\sum y^2$	Anal. mean square	Treatments plus error $\sum x_m^2$	$\sum x_1x_2$	$\sum x_1y$	$\sum x_2y$	$\sum y^2$	$\sum y^2 - \sum_m \hat{\beta}_m(\sum x_m y)$	Adj. $\sum y^2$	Mean square
Main plot																
A	1	32	−32	−245	245	1876	5832	5832**	7(32)	−32	761	129	7845	5004	4376	4376**
Error (W)	6	6(32)	0	1006	−116	5969	628	105							628	$105 = s_W^2$
Subplot																
B	1	32	−32	−151	151	713	694	694								
C	1	32	32	271	271	2295	314	314								
BC	1	32	−32	17	−17	9	11	11								
AB	1	32	32	−259	−259	2096	5751	5751**	19(32)	32	1255	1289	14081	9024	5178	5178**
AC	1	32	−32	−165	165	851	830	830								
ABC	1	32	32	197	197	1213	23	23								
Error (E)	18	18(32)	0	1514	1548	11985	3846	240							3846	$240 = s_E^2$

Missing value and $\hat{\beta}_m$ estimates:

Main plot: The $\hat{\beta}_{m_W}$ estimates satisfy: $\begin{cases} 6(32)\,\hat{\beta}_{1_W} + 0\,\hat{\beta}_{2_W} = 1006. \\ 0\,\hat{\beta}_{1_W} + 6(32)\hat{\beta}_2 w = -116. \end{cases}$ Therefore $\begin{cases} \hat{\beta}_{1_W} = 5.24 \text{ and} \\ \hat{\beta}_{2_W} = -0.60 \end{cases}$ $\begin{cases} \text{Missing value}_{1_W} = 168 \\ \text{Missing value}_{2_W} = -19 \end{cases}$

Subplot: The $\hat{\beta}_{m_E}$ estimates satisfy: $\begin{cases} 18(32)\,\hat{\beta}_{1_E} + 0\,\hat{\beta}_{2_E} = 1514. \\ 0\,\hat{\beta}_{1_E} + 18(32)\,\hat{\beta}_{2_E} = 1548. \end{cases}$ Therefore $\begin{cases} \hat{\beta}_{1_E} = 2.63 \text{ and} \\ \hat{\beta}_{2_E} = 2.68 \end{cases}$ $\begin{cases} \text{Missing value}_{1_E} = 84 \\ \text{Missing value}_{2_E} = 86 \end{cases}$

Subtraction of the correct error term, Error (W) = 628, from this gives the A sum of squares, 4376, which will provide an exact test for treatment effect A.

. . . . . .

The author is greatly indebted to O. Kempthorne for the basic theorems on which the technique is based and for suggestions on the preparation of this paper.

REFERENCES

Anderson, R. L. [1946] Missing plot techniques. *Biometrics, 2:* 41–47.

Bartlett, M. S. [1937] Some examples of statistical methods of research in agriculture and applied biology. *J. Roy. Stat. Soc. Suppl., 4:* 137–170.

Hartley, H. O. [1956] Programming analysis of variance for general purpose computers. *Biometrics, 12:* 110–122.

Kempthorne, Oscar [1952] *The design and analysis of experiments.* John Wiley and Sons, Inc., New York.

Nelder, J. A. [1954] A note on missing plot values. *Biometrics, 10:* 400–401.

Norton, H. W. [1955] A further note on missing data. *Biometrics, 11:* 110.

Smith, H. F. [1957] Missing plot estimates. *Biometrics, 13:* 115–118.

Truett, J. T. and Smith, H. F. [1956] Adjustment by covariance and consequent tests of significance in split-plot experiment. *Biometrics, 12:* 23–39.

ABSTRACTS

Papers read at the 4th Biometric Colloquy at Bad Nauheim, Germany, January 25–27, 1957

416 A. AUGSBERGER (Nürnberg). **Statistical Tests on Therapeutic Trials of Chronic Diseases, Performed on the Estimation of an End Value.**

Therapeutic trials in chronic diseases involve observations of quantitative (e.g. blood pressure in hypertension) or graded symptoms in the same patient during a first period (where there may be no therapy or placebo or drug A) and during a second period (drug B). The physician's main question is whether the value of the symptom at the end of the second period is more favorable than one would expect had the first period continued unaltered. Fitting a curve to the observed values (or their logarithms, resp.) of each period (a straight line as a first approximation) the author examines the difference between the value of the symptom estimated by extrapolation of the first period on the last day of the second period and either the observed or the regression value of the second period on the same day. The standard error of the difference is obtained from the pooled error variance about the two regression lines.

417 R. K. BAUER (Krefeld). **Statistical Investigations with Automatic Computers.**

Automatic computers used for investigations on statistical data (a) reduce writing by hand to a minimum, and computing by hand or table machines to zero; (b) eliminate nearly all copying mistakes; (c) replace a slow staff of arithmeticians; (d) warrant high actuality of computing results by high speed of computing procedure; (e) render possible a more profound analysis of data. Therefore not only the largest, but all statistical investigations of an institute, a hospital or of industrial works may be done by automatic computers.—To get these advantages three conditions have to be fulfilled: (i) an efficient and reliable computing aggregate has to be on hand without delay; (ii) data have to be surveyed on punchable blanks, their documentation has to

be made on punch cards; (iii) a library of universal computing programms for all standard statistical methods has to be compiled most quickly.— (i) was discussed; on (ii) some illustrations were shown; on (iii) a Bell-programmed total correlation and regression analysis by means of the IBM Magnetic-Drum-Data-Processing-Machine type 650 was demonstrated.

418 H. GRIMM (Jena). **Some Problems of Bacterial Distribution and Counting.**

1) In counting particles, time may be saved when a modified counting method is employed. The mean and standard error are estimated from the truncated Poisson. A little loss in precision is compensated by an essential saving of time.

2) The type of distribution is readily recognized by plotting on Poisson cumulative probability paper. The various cumulative curves (Poisson, binomial, Neyman, negative binomial, log-series) become evident from this diagram, and their parameters may be estimated by means of transparent stencils. Some empirical compound distributions are discussed.

3) Evidence is given of the difficulties involved in photoelectric counting.

419 P. KÜHNE (Berlin-Charlottenbg.).

1) Provisional statistical estimation of the significance of differences between means reported without variance.

In medical papers not uncommonly mean values are reported without simultaneously giving any information on the variance. By a confidence procedure—with N known—a limit can be estimated for the standard deviation beyond which a difference between means would have to be regarded as insignificant. A comparison with standard deviations, recorded elsewhere, of appropriate experimental values then permits inferences on differences of means unsatisfactorily reported. If the significance of a difference between the mean of a "pathological" group and that of a normal one has to be tested, the 's' limit of the pathological group can be isolated since the parameters of most clinically normal values can be found in tables.

2) The use of the inverse sine transformation for adequate representation of results expressed as ratios.

Tests of significance of differences between series of ratios from the same experimental material by means of the inverse sine transformation can be reduced to a single calculation, since the standard deviation of

the difference depends on the total number of observations only. An empirical solution for the necessary correction for small N by entering the t-distribution with $(N_1 + N_2)/2$ d.f. is given and demonstrated on a clinical example. Furthermore a solution is given for the problem of how many samples have to be taken for the detection of contamination of low incidence (bacterial contamination in industrial production of ampoulated remedies).

3) A method for estimating the significance of rank correlations.

In testing the correlation between continuous and coarsely grouped variates a "genuine" distribution for uncoordinated behavior depending solely on the total number can be calculated. This null-hypothesis of the difference of rank-numbers can be tested by chi-square against the sample distribution of experimental results. The calculable half-value of the distribution is chosen for chi-square grouping.

4) The distribution of "pathological" values.

Assuming the total population of any biological variate as distributed normally the values of morbid individuals should be included within this total distribution as a tail of extremes. According to clinical usage those "pathological" values are separated from those of the "normal" population. This corresponds to a subtraction of two normal distributions differing in their σ. The resulting theoretical distribution is markedly asymmetrical and is compared with empirical distributions encountered in clinical research. It appears that only values collected under policlinical or ambulatory conditions conform to theory whereas ward groups tend to form a symmetrical distribution.

420 W. LUDWIG (Heidelberg). **On Some Partially Solved or Unsolved Problems in Biomathematics.**

A short review on problems attacked together with co-workers: (1) Animals belonging to a certain taxonomic group are consuming energy, by means of an organ I proportional to L^2 (L = body length), by means of organ II proportional to L^3; what follows for the energy + L^2—relation in a log-log graph? Experimental verification by insect larvae and other animals.—(2) How many bacterial spores may x ml Agar culture-medium spread on a Petri dish of diameter y contain on the average, such that a culture of diameter z can be said to be originated from one single spore with a probability of 99% and over (WETTE)?—(3) Rate of progress of annidation (cf. LUDWIG 1950; WARTMANN unpubl.).—(4) Territory problems. I. Coins falling on a desk successively; any coin overlapping one already deposited will be removed. Occupation density maximum?—(5) Mimicry: Proportion of indifferent animals to models to imitators; e.g. coral snakes of America.—(6) Dem-

onstration of lack of adequate statistical methods in certain "regression" problems in biology.—

421 G. OBERHOFFER (Bonn). **On Some Principles of Therapeutic Trials.** (Construction of homogeneous material for comparative trials.)

The therapeutic trial should comply with the requirements resultant from the standards of an exact experimentation without violating the patient's faith. The therapeutic trial always has to be based on trial and control groups which have to be homogeneous. Discussion of some important principles set up by *Martini* for the performance of therapeutic trials: with acute diseases comparison of groups consisting of different patients, with chronic diseases comparison of different periods of the same patient.

422 K. RIEGEL (Hamburg). **On the Problem of Scaling and the Factor Analysis of Psychological Tests.**

So far test theory has yielded only little information on the problem of exact scaling methods as promoted in the field of experimental psychology, e.g. by *Thurstone, Gulliksen, Guilford, Coombs, Stephens,* and others. Generally three ways have been used to transfer exact scaling methods into test theory, which may be called (1) the "realistic" (2) the "instrumentalistic", and (3) the method of "item validation". For demonstration purposes test data for a group of persons from the standardization sample of the *Hamburg-Wechsler* intelligence test were weighted anew in such a way as to express the sociological meaning of passing the items by new scores in order to improve the scaling base of the test. Necessarily, the reliability of the test and the subtests, and accordingly the usefulness of the test in a technical sense, were decreased by using the new scoring method. Correlations between new and old scores point out the danger of factor-analyzing test data of a similar kind and emphasise the need to limit factor analysis to item statistics.

423 W. SEYFFERT (Berlin). **A Method of Estimation of Interallelic Interaction, a Contribution to the Gene Dosage Problem.**

In estimating the valence of the single alleles of a pair of allelomorphs the measurements of the different diploid and tetraploid genotypes which have a constant genetic background consist of the gene dosage and the genome dosage. The influence of each of these contributions

has been evaluated by the method of least squares, yielding three equations of the eighth degree. Using *Newton-Raphson's* iteration these result in the following estimates: x = valence of the dominant allele, y = valence of the recessive allele, z = valence of the genetic background. (To appear in Z. Vererbgsl. *88*; 1957.)

424 F. STALLMAN (Giessen). **Mathematical Notes on Electrocardiography.**

The scope of electrocardiography is the inference of the distribution of electric elements in the heart from measurements of body surface potentials. By means of a theorem of *Helmholtz'*, based on the symmetry of the *Green* function, the problem can be reduced to the construction of presupposed fields of electric currents in the body interior. It can be shown that the elements of the heart constitute a single heart dipole, the position, direction, and magnitude of which can be determined from potential measurements. The mathematical problems concerning the development of this theory and methods of determination are outlined. (To appear in: Arch. Kreislaufforsch.)

425 F. SÜLLWOLD (Göttingen). **The Concept of "Simple-structure" and the Problem of Factorial Invariance.**

It is demonstrated that factors extracted by the centroid method are to be considered as an arranging means only and that they are of no concrete value. The necessity is pointed out for a substitution of the voluntary original reference system. Discussion of the problem of correct or adequate position, respectively, of coordinates in view of the manifold of rotations, with regard to different criticisms concerning the same heuristic principle.

426 E. WALTER (Göttingen). **Some Simple Unbiased Nonparametric Tests for Symmetry with Respect to Zero.**

Given n differences $x_i = y_{1i} - y_{2i}$. A class of strictly unbiased nonparametric rank order tests is considered for testing the hypothesis $f(x) = f(-x)$, i.e. the symmetry of the probability density $f(x)$ with respect to zero, under the assumption of absolute continuity of the distribution function $F(x)$. This class contains amongst others (a) a modified sign test (in contrast to the sign test by *R. A. Fisher* the absolute smallest difference is discarded) (b) a test of size $\alpha = 2^{-k+1}$, for which the hypothesis $f(x) = f(-x)$ has to be rejected when the k absolute largest differences have the same signs. This test is the optimum order test against distributions which are truncated but originally

symmetrical with respect to zero. Furthermore, this test has higher efficiency than the sign test when the distribution is normal and n small.

Papers presented at joint I.M.S. and Biometric Society (ENAR) Sessions, Washington, March 7–9, 1957.

VACLAV EDVARD BENES (Bell Telephone Laboratories, Murray Hill, N. J., U.S.A.). **427 The Joint Distribution of a Set of Sufficient Statistics for the Parameters of a Simple Telephone Exchange Model.**

This paper considers a simple telephone exchange model which has an infinite number of trunks and in which the traffic depends on two parameters, the calling-rate and the mean holding-time. It is desired to estimate these parameters by observing the model continuously during a finite interval, and noting the calling-time and hang-up time of each call, insofar as these times fall within the interval. It is shown that the resulting information may, for the purpose of this estimate, be reduced without loss to four statistics. These statistics are the number of calls found at the start of observation, the number of calls arriving during observation, the number of calls leaving during observation, and the average number of calls existing during the interval of observation. The joint distribution of these sufficient statistics is determined (in principle) by deriving a generating function for it. From this generating function the means, variances, covariances, and correlation coefficients are obtained. Various extimators for the parameters of the model are compared, and some of their distributions, means, and variances presented.

R. C. BOSE (Department of Statistics, University of North Carolina, Chapel Hill, N. C., U.S.A.). **428 On a Problem in Abelian Groups and the Construction of Fractionally Replicated Designs.**

Consider an Abelian group of order s^n, generated by n letters $A_1, A_2, \cdots, A_n$ with the relations $A_1^{s_1} A_2^{s_2} \cdots A_n^{s_n} = I$, where I is the identity and s is a prime. If $G = A_1^{x_1} A_2^{x_2} \cdots A_n^{x_n}$ is any element of the group, then the number of non-zero exponents x_i may be called the length of G. Given an integer $r < n$, the problem is to find a subgroup of order s^r, generated by r independent elements $G_i = A_1^{x_{i1}} A_2^{x_{i2}} \cdots A_n^{x_{in}}$ such that the minimum length of the elements in the subgroup (except the length of the unit element) is greater than or equal to k. Consider the finite projective space $PG(r - 1, s)$. To any point $x = (x_1, x_2, \cdots, x_r)$ of this space, assign a non-negative integer m, which may be considered the measure of x, in such a way that the total measure for the space is n. To a point of measure m

associate m different letters chosen out of A_1 , A_2 , $\cdots$, A_n , each of these letters being assigned to one and only one point. Let $G_i = A_1^{x_{i1}} A_2^{x_{i2}} \cdots A_n^{x_{in}}$ where x_{ij} is the ith coordinate of the point to which A_j is associated. It is proved that the length of the element $G_1^{\lambda_1} G_2^{\lambda_2} \cdots G_r^{\lambda_r}$ is the measure of the set of points not lying on the linear space $\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_r x_r = 0$. For example let $n = 10$, $r = 4$, $s = 3$. We can find exactly 10 points on the unruled quadric $2x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4 = 0$, in $PG(3, 3)$. Any plane meets this quadric in four points, or in one point. If we assign the measure one to each point of the quadric and associate with it one of the letters A_i we get $G_1 = A_2A_3A_4A_5A_6A_7$, $G_2 = A_1A_3A_4A_5^2A_6^2A_8$, $G_3 = A_1A_2A_4A_5A_6A_9$, $G_4 = A_1A_2A_3A_5^2A_6^2A_{10}$. The length of the words (other than the identity) of the subgroup generated by G_1 , G_2 , G_3 , G_4 must be either 6 or 9. If we take this subgroup as the fundamental identity for generating a $1/2^4$ fraction in a factorial design with 10 factors, then all the aliases of a main effect will have five or more factors, and all the aliases of a two factor interaction will have four or more factors.

429 BRADLEY BUCHER (Princeton University, N. J., U.S.A.). **The Recovery of Intervariety Information.**

Assume, in the incomplete block model, $y_{ij} = m + b_i + v_j + s_{ij}$, that the block effects are independently distributed with mean 0 and variance β^2, the error terms e_{ij} are independently distributed with mean—and variance α^2, and that the variety effects t_1 , $\cdots$, t_k , are fixed effects and that t_{k+1} , $\cdots$, t_n , are independently distributed with mean 0 and variance γ^2 . . Then in estimating any linear combination of the variety effects, say $a_1t_1 + a_2t_2 + \cdots + a_kt_k$, we may make use of information among the varieties t_{k+1} , $\cdots$, t_n . Minimum variance linear unbiased estimates are obtained for such combinations for a large class of incomplete block designs. In general, these estimates have smaller variance than analogous estimates obtained using only inter- and intra-block recovery. For balanced incomplete blocks the estimate with inter-variety recovery is shown to be the same as the combined intra- and inter-block estimate. Several techniques are developed which are useful for finding estimates using intervariety recovery. The problem of estimating γ^2 is discussed. Useful applications of the technique of intervariety recovery are considered.

430 J. T. CHU AND F. C. LEONE AND C. W. TOPP (Case Institute of Technology and Fenn College, Ohio, U.S.A.). **Some Uses of Quasi-Ranges II.**

In "Some uses of quasi-ranges" (*Ann. Math. Stat. 28*, 1) methods

are given of using quasi-ranges to obtain confidence intervals for, and tests of hypotheses about, some measures of dispersion of a given distribution, (such as the interquantile distance and the standard deviation). In this paper, further research is done on the selection of quasi-ranges for making inferences about the standard deviations of the normal, rectangular, and exponential distributions. The methods are also extended to the coefficient of variation, the difference and ratio of interquantile distances and standard deviations of two given distributions, etc. Tables are given to facilitate applications.

431 A. BRUCE CLARKE (University of Michigan, Ann Arbor, Michigan, U.S.A.) **Maximum Likelihood Estimates in a Simple Queue.**

A simple stationary queuing process is a queuing process having a Poisson input (with parameter λ), and a negative exponential service time (with mean $1/\mu$, $\mu > \lambda$). Let ν = the initial queue size, x_i = the time of the ith arrival, y_i = the "busy time" up to the ith departure. The sequences $\{x_i\}$ and $\{y_i\}$ then represent the transition times of *independent* Poisson processes (parameters λ and μ), and $\{\nu\}$, $\{x_i\}$, $\{y_i\}$ together characterize the process. By observing the process for a fixed "busy time" τ and using the above comment, maximum likelihood estimates for λ and μ may be obtained in terms of ν, m = the total number of departures, T = the time of the mth departure, and n = the total number of arrivals up to time T. Under certain conditions these estimates of λ and μ may be approximated by $(n + \nu)/T$ and $(m - \nu)/\tau$.

432 EARL DIAMOND (University of North Carolina, U.S.A.). **Extension of Some Results Given by Mitra on "Statistical Analysis of Categorical Data".**

This is a follow up of two previous papers [(1) "Some non-parametric generalizations of analysis of variance and multivariate analysis" by S. N. Roy and S. K. Mitra, *Biometrika*, December, 1956 and (2) "Contributions to the statistical analysis of categorical data" by S. K. Mitra, North Carolina Institute of Statistics Mimeograph Series No. 142]. We start from a product of multinomial distributions of the form $\phi = \prod_j [n_{0j}! \prod p_{ij}^{n_{ij}} / \prod_i n_{ij}!]$ with $\sum_i p_{ij} = 1, i = i_1, i_2, \cdots, i_k$; $j = j_1, j_2 \cdots j_l$; $i_1 = 1, 2, \cdots, r_1$; $\cdots$; $i_k = 1, 2, \cdots, r_k$; $j_1 \varepsilon (s_1)_{i_2 \cdots i_l}$ (a subset of s_1 depending on the subscript set $j_2 \cdots j_l$); $j_2 \varepsilon (s_2)_{i_3 \cdots i_l}$; $\cdots$; $j_{l-1} \varepsilon (s_{l-1})_{i_l}$ and $j_l = 1, 2, \cdots, s_l$. We next consider two hypotheses $H_0^{(1)} : p_{ij} = f_{ij}^{(1)} (\theta_1, \cdots, \theta_{t_1})$ subject to $g_m^{(1)}(\theta_1, \cdots, \theta_{t_1}) = 0$ ($m =$

$1, 2, \cdots, u_1 < t_1$) and $H_0^{(2)} : p_{ij} = f_{ij}^{(2)}(\theta_1', \cdots, \theta_{t_2}')$ subject to $g_m^{(2)}(\theta_1', \cdots, \theta_{t_2}') = 0 (m = 1, 2, \cdots, u_2 < t_2)$, $t_1, t_2 <$ total number of cells $-$ total number of multinomial distributions. Each hypothesis is a composite one in which the θ or θ' are the nuisance parameters and $f_{ij}^{(1)}, g_m^{(1)}, f_{ij}^{(2)}$ and $g_m^{(2)}$ are known functions. Tests are taken over from references (1) and (2), and the asymptotic powers of the tests and the conditions for asymptotic independence are derived which are extensions of similar conditions for more special cases discussed in (2).

433 JOHN J. GART (Virginia Polytechnic Institute, Blacksburg, Virginia, U.S.A.). **An Extension of the Cramér-Rao Inequality.**

Consider a frequency function $f(x \mid \boldsymbol{\theta})$ where $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_s)$, the function being specified when $\boldsymbol{\theta}$ is specified. The parameter $\boldsymbol{\theta}$ has a density $g(\boldsymbol{\theta})$ independent of x. Let $\mathbf{X} = (x_1, x_2, \cdots, x_n)$ be a random sample from a randomly chosen population having the specified frequency function. Then if $\phi = \prod_{i=1}^{n} f(x_i \mid \boldsymbol{\theta})$ and t_k (independent of $\boldsymbol{\theta}$) is an estimate of θ_k, $1 \leq k \leq s$, there follows a form similar to the Cramér-Rao Inequality,

$$EE[(t_k - \theta_k)^2 \mid \boldsymbol{\theta}] \geq \{E[E(t_k \mid \boldsymbol{\theta}) - \theta_k]\}^2 + E^2\left(\frac{\partial E(t_k \mid \boldsymbol{\theta})}{\partial \theta_k}\right)\left\{EE\left[\left(\frac{\partial \ln \phi}{\partial \theta_k}\right)^2 \,\middle|\, \boldsymbol{\theta}\right]\right\}^{-1}$$

The equality is reached if and only if t_k is an unbiased sufficient statistic having the normal distribution with constant variance. In this case the equality holds regardless of the form of $g(\boldsymbol{\theta})$.

434 SHANTI S. GUPTA AND MILTON SOBEL (Bell Telephone Laboratories, Allentown, Pennsylvania, U.S.A.). **On Selecting a Subset which Contains All Populations Better Than a Standard.**

Populations $\prod_i$ ($i = 0, 1, \cdots, p$) are given with a common Koopman-Darmois distribution of known form differing only in the value of the unknown parameter $\tau_i (i = 1, 2, \cdots, p)$; cases of known and unknown (associated with the standard $\prod_0$) are treated separately. Location and scale parameter problems are both treated. In some problems $\prod_i$ is defined as better than $\prod_0$ if $\tau_i > \tau_0$; in others if $\tau_i < \tau_0$. A procedure is given in each case for selecting a small subset so that, for any true configuration, the probability of including all $\prod_i$ equal to or better than $\prod_0$ is at least P^*, $P^* < 1$ being preassigned. For the location parameter, with unknown, the procedure is to retain all $\prod_i$ with $\bar{w}_i = \sum_{j=1}^{n_i} w(x_{ij}) \geq \bar{w}_0 - d/\sqrt{n_i}$; here $\bar{w}_i$ is sufficient for

$\tau_i (i = 0, 1, \cdots, p)$. For scale parameter problems, with smaller more preferable, the procedure retains all $\prod_i$ with $\sum_{j=1}^{n_i} w(x_{ij}) \leq (1 + d) \sum_{j=1}^{n_0} w(x_{0j})$. In several problems the value of d is computed and tables are given for different P^* and p-values; in others transformations are used to "normalize" the problem. The normal and chi-square distributions are used as applications. Problems involving binomial and Poisson distributions are treated separately with and without normalizing transformations.

435 JOHN GURLAND (Iowa State College, Ames, Iowa, U.S.A.). **Generalized Quantal Response in Biological Assay.**

The quantal (all-or-none) response in biological assay refers to a response in which one of two possible outcomes occurs. In a bioassay such as that of an insecticide based on mortality of the housefly, say, there are however three possible outcomes, namely alive, moribund, dead. The present paper considers a generalized quantal response in which two or more outcomes are possible. Whether one uses normits (cf. probits) or logits or other transformations, a general method of analyzing the data is developed which makes explicit use of all the possible outcomes and hence is more efficient than the common procedure of pooling some outcomes (for example moribund and dead) in order to make the response all or none. Further, a technique analogous to that used in discriminant functions is suggested as a method which makes more efficient use of the data than the pooling method mentioned above.

436 ANDIE LUBIN (Walter Reed Army Institute of Research, Washington, D. C., U.S.A.). **A Rank Order Test for Trend in Correlated Means.**

In many experiments the major interest is not in the amount of difference caused by the treatments but the rank-order which results. This is especially true when successive measurements are made on the same subject, and the "treatments" are simply varying amounts of fatigue, sleep loss, etc., i.e., some function of time. For such studies the null hypothesis is that no trend exists and generally the only alternative hypothesis is a rank-order that can be specified by the experimenter.

A. R. Jonckheere (1954, A distribution-free k sample test against alternatives, 41, *Biometrika*) has used Kendall's tau to obtain a general statistic, P, for testing the agreement between a hypothesized rank-order for n objects or scores and a set of observed rankings of the n

scores by m judges. From this general approach, he derives a test for trend as a special case.

As an alternative to Jonckheere's P, a statistic J based on Spearman's S (d^2) is examined. It is the sum of the $S(d^2)$ values computed between the n observed rankings of the n scores and the hypothesized rank-order of the n scores. K, the average rank order correlation between the m rankings and the hypothesized rank-order, a simple algebraic function of J.

It is shown that J is slightly more sensitive than Jonckheere's P statistic for small values of n, but that P tends to normality faster than J.

437 J. A. McFADDEN (U. S. Naval Ordnance Lab., Silver Spring, Maryland, U.S.A.). **The Variance of Zero-Crossing Intervals.**

Two expressions are given for the variance of the intervals between successive zeros of a random process. It is assumed that the successive intervals form a Markoff chain. If $x(t)$ is a random process, let $y(t) = 1$ when $x(t) \geq 0$ and $y(t) = -1$ when $x(t) < 0$. Let β be the expected number of zeros per second and let x be the correlation coefficient between two successive zero-crossing intervals. Then the variance is $\sigma^2 = (2A/\beta)(1 + x)/(1 - x)$, or alternatively, $\sigma^2 = [(1 + 2B)/\beta^2]$ $(1 - x)/(1 + x)$ where $A = \int_0^\infty r(\tau)\, d\tau$ and $B = \int_0^\infty [Q(\tau) - \beta]\, d\tau$. $r(\tau)$ is the autocorrelation function of the process $y(t)$ and $Q(\tau)\, d\tau$ is the conditional probability of a zero between $t + \tau$ and $t + \tau + d\tau$, given a zero at time t.

438 JUNJIRO OGAWA (University of North Carolina, North Carolina, U.S.A.). **A Limit Theorem of Cramér and Its Generalization.**

As a generalization of Doob's theorem, H. Cramér states the following. Theorem: Suppose we have for every $\nu = 1, 2, \cdots$, $y_\nu = Ax_\nu + z_\nu$, where x_ν, y_ν and z_ν are n-dimensional random variables, while A is a matrix of order $n \times n$ with constant elements. Suppose further that as $\nu \to \infty$, the n-dimensional distribution of x_ν tends to a certain limiting distribution, while z_ν converges in probability to zero. Then y_ν has the limiting distribution defined by the linear transformation $y = Ax$ where x has the limiting distribution of the x_ν (H. Cramér: *Mathematical Methods of Statistics*, Princeton, 1946, pp. 299-300). Cramér skips the proof of this theorem. In this paper, the complete proof of this theorem will be given and two theorems which are generalizations of this theorem and are useful in statistics will be proved.

439 JUNJIRO OGAWA (University of North Carolina, North Carolina). **On the Mathematical Principles Underlying the Theory of the χ^2 Test.**

The rigorous proof of the theorem that the χ^2 statistic has the limiting chi-square distribution with degrees of freedom reduced by the number of the independent parameters which were estimated from the frequency data, was first given by H. Cramér in his famous book *Mathematical Methods of Statistics*, Princeton [1946], but some steps of the proof were skipped. Later on S. N. Roy and S. K. Mitra (*Biometrika*, 43 [1956]) and S. K. Mitra (Thesis, Univ. of N. C., 1956, reasoned along the same lines and got theorems adjusted to various physical situations. The purposes of this paper are to present a complete and self-contained proof of Cramér's theorem on the one hand, and on the other to explain how the proof of the related theorems got by S. N. Roy and S. K. Mitra could be thrown back on that of Cramér's theorem from the mathematical point of view.

440 JUNJIRO OGAWA (University of North Carolina, North Carolina, U.S.A.). **A Further Contribution to the Theory of Systematic Statistics.**

Up to 1945 the main interest of statistical estimation has been in the "efficient estimator," but from the point of view of practical use, it seems reasonable to inquire whether comparable results could have been obtained by a smaller expenditure. F. Mosteller [1946] proposed the use of systematic statistics in this connection. The author [1952] developed a systematic theory of estimation and testing hypothesis with respect to the location and scale parameter of a population whose density depends on only these two parameters.

There are many cases in which the samples are by their very nature ordered in magnitude, for example in a life test of electric lamps. In such cases the population probability distributions are usually supposed to be exponential. Thus, at least for the exponential distribution, estimation and testing of a hypothesis based upon systematic statistics are of great importance from the standpoint of practical application.

There will be presented in this paper the table of the optimum spacings of the selected sample quantiles, corresponding best estimators, and a discussion on the testing procedure of a statistical hypothesis on the scale parameter of the exponential distribution $f(x) = (1/\sigma)e^{-(x/\sigma)}$ for $x > 0$.

441 DAVID ROSENBLATT (American University, Virginia, U.S.A.). **On the Stochastic Structure of Minkowski-Leontief Systems.**

A linear system $x(I - A) = w$ is said to be of *Minkowski-Leontief type* if A is a finite non-negative square matrix of order n with no row sum exceeding unity and x, w are non-negative row vectors. A non-null solution x of such a system is called *admissible*. *Theorem*: *Every system of Minkowski-Leontief type* $x(I - A) = w$ *which exhibits at least one admissible solution is equivalent to a unique system* $\tilde{x}(I - \tilde{A}) = \theta$, *where* $\tilde{A}$ *is a stochastic matrix depending on* A *and* w *and* θ *is a null vector of dimension at most* $n + 1$. *Every admissible solution of* $x(I - A) = w$ *(appropriately extended or contracted) is proportional to a convex linear combination of the stationary stochastic vectors of* $\tilde{A}$. If A is non-stochastic, $w \neq \theta_n$, let $\tilde{A}$ denote the matrix $\begin{Vmatrix} A & b \\ w^* & 0 \end{Vmatrix}$ where $b = (I - A)e'$, $w^* = \lambda_w^{-1} w$, $\lambda_w = we'$, and e is the row vector with all elements unity. If $(I - A)^{-1}$ exists and $w \neq \theta_n$, there exists a single ergodic set of indices; if w is positive the stationary vector of $\tilde{A}$ is positive. Clearly, $\tilde{x} = (w(I - A)^{-1}, \lambda_w)$. If $w = \theta_n$ and $(I - A)$ is singular, $\tilde{A}$ is taken as $\begin{Vmatrix} A_r & 0 \\ 0 & I_{n-r} \end{Vmatrix}$ where A_r is the largest stochastic principal sub-matrix of A. Systems of the present type occur in economic input-output analysis and generally in socio-physical models based on "balanced-margin" tables, i.e., non-negative square matrices X such that $cX = cX'$.

442 DAVID ROSENBLATT (American University, Virginia, U.S.A.). **On the Stochastic Structure of Minkowski-Leontief Systems, II.**

Consider a system $x(I - A) = w$ of Minkowski-Leontief type such that $(I - A)^{-1}$ exists. Clearly, $(I - A)^{-1}$ exists if and only if A contains no stochastic principal sub-matrix. In a static economic input-output context the element a_{ij} is designated as the input (per unit output) to industry or activity i procured from industry j; w_j, x_j are respectively final output and total output (or activity level) of the jth industry. Consider the uniquely corresponding system $\tilde{x}(I - \tilde{A}) = \theta$, where $\tilde{A}$ is stochastic. The unique stationary stochastic vector of $\tilde{A}$ is given by $(p_{n+1}w^*(I - A)^{-1}, p_{n+1})$. The "multiplier" $\mu = \sum_{j=1}^{n+1} x_j/\lambda$ is given by $1/p_{n+1}$ where $\lambda = x_{n+1} = we'$. Given a non-singular matrix $(I - A)$, the following relation holds in components of an admissible solution for any w: $\sum_{j=1}^{n} (1 - r_j)x_j - x_{n+1} = 0$, where $x_{n+1} = we'$ and r_j is the jth row sum in A. The latter relation is the technical production-

possibility function of the economy in an input-output sense; $-\Delta x_j/\Delta x_k = (1 - r_k)/(1 - r_j)$, $\Delta x_j/\Delta x_{n+1} = 1/(1 - r_j)$, $j \neq k$; $j, k = 1, \cdots, n$, are the invariant "substitution ratios" of the system, obviously independent of w. Let x be an admissible solution of $x(I - A) = w$, $(I - A)$ singular or not, and let $D(x, \lambda)$ be a diagonal matrix with components of x and λ on the diagonal. Then $D(x, \lambda)\tilde{A}$ is a "balanced-margin" table. Consistent with a noted "substitution" result, $\sum_{j=1}^{n} K_j w_j = x_{n+1} = \lambda$ where $K_j = 1$ for all j independently of w.

443 DAVID ROSENBLATT (American University, Virginia, U.S.A.). **On the Stochastic Structure of Minkowski-Leontief Systems, III.**

Consider any system $x(I - A) = w$ of *M-L* type. The following "aggregation" problem is of interest. Let an *aggregation matrix* C be an $n \times r$ stochastic matrix of incidence type, $1 \leq r < n$. Let $B = f(A)$ be a *M-L* matrix of order r. We consider conditions under which $\hat{x}AC = \hat{x}CB$ obtains for admissible solutions $\hat{x}$ of a system $x(I - A) = w$. The following case is of special interest. Let a *weight matrix* E be an $n \times n$ diagonal matrix with non-negative entries on the principal diagonal. A *consolidation* of a matrix A of *M-L* type is an $r \times r$ matrix $B \equiv B(A; C, E) = (C'EC)^{-1}C'EAC$, $1 \leq r < n$. "Faithful consolidation" of a stochastic system $x(I - A) = \theta$ is characterized from the standpoint of ergodic structure; the condition $AC = CB(A; C, E)$ is of particular interest. A general consolidation condition for *M-L* systems is related to the "combining-of-classes" condition of stochastic learning theory. The following is of economic interest: the existence of $(I - B)^{-1}$ does not in general imply the existence of $(I - A)^{-1}$, and conversely. In the static input-output model of II, the ergodic structure of $\tilde{A}$ of the equivalent system (and the role of mean recurrence time $1/(p_{n+1})$) suggest that the stationary stochastic vector $\tilde{y}$ of $\tilde{A}$ be computed iteratively using successive powers of $\tilde{A}$, yielding $\tilde{x}$, in lieu of matrix inversion with or without consolidation; in most applications, $\lim_{k\to\infty}\tilde{A}^k$ exists.

444 S. N. ROY AND M. D. MOUSTAFA (University of North Carolina, North Carolina, U.S.A.). **Testing of Hypotheses on a Mixture of Variates Some of Which Are Continuous and the Rest Categorical.**

We start from a $k + l$-variate distribution in which k variates are continuous and l variates are categorical. The k variates are assumed to have a conditional multivariate normal distribution with respect to the l categorical variates which are assumed to have a multinomial

distribution. Appropriate hypotheses are framed in this situation, analogous to the customary hypotheses on a single multivariate normal distribution (or to those in references (1) and (2) of the previous abstract), large sample tests of such hypotheses are developed and some of their properties studied. Next, instead of assuming a single multinomial distribution on the l categorical variates, a product of multinomial distributions is assumed and hypotheses are framed in this situation analogous to the customary ones for several multivariate normal distributions or to those in references (1) and (2), and large sample tests of such hypotheses and some of their properties are studied.

445 S. N. ROY AND R. GNANADESHIKAN (University of North Carolina, North Carolina, U.S.A.). **Confidence Bounds Associated with Multivariate Analysis of Variance.**

We start from the same set-up as in the previous paper. The S^* and S (to be called respectively the dispersion matrix "due to the hypothesis" and the dispersion matrix "due to the error") are the exact analogs of the variance "due to the hypothesis" and that "due to the error" in the customary univariate analysis of variance. Given any level α, we can pick up a constant c_α from the tables mentioned in the previous paper and make, with a probability greater than or equal to $1 - \alpha$, the confidence interval statement: $c_{\max}^{1/2}(sS^*) - [sc_\alpha]^{1/2} \times c_{\max}^{1/2}(S) \leq c_{\max}^{1/2}[\eta' U \eta] \leq c_{\max}^{1/2}(sS^*) + [sc_\alpha]^{1/2} c_{\max}^{1/2}(S)$, where $U(s \times s)$ is a non singular matrix given (in the paper) in terms of A and C, and $c_{\max}^{1/2}[\eta' U \eta]$ is zero if and only if $\eta = 0$, i.e., H_0 is true. With a joint probability greater than or equal to $1 - \alpha$ we can also make simultaneous confidence interval statements including the one given above and others exactly similar to this but in terms of $S^{(i)}$, $S^{(i)*}$, $\eta^{(i)}$ (for $i = 1, 2, \cdots, p$) and next terms of $S^{(i,j)}$, $S^{(i,j)*}$, $\eta^{(i,j)}$ (for $i \neq j = 1, 2, \cdots, p$) and so on, where $S^{(i)}$ and $S^{(i)*}$ stand respectively for truncated matrices after cutting out the ith row and ith column from S and S^*, $\eta^{(i)}$ for η with the ith column cut out, $S^{(i,j)}$, $S^{(i,j)*}$ for S and S^* with the ith and jth rows and columns cut out, $\eta^{(i,j)}$ for the ith and jth columns cut out, and so on.

446 S. N. ROY (University of North Carolina, North Carolina, U.S.A.). **Multivariate Analysis of Variance.**

Consider a model under which we have stochastic variates $X(p \times n) = [x_1 \cdots x_n]p$ such that the x_i (for $i = 1, 2, \cdots, n$) are independent $N[E(x_i), \Sigma]$, $E(X') = A(n \times m) \times \xi(m \times p)$, A (to be called the

design matrix) is a matrix of constants given by the design of the experiment, ξ is a matrix of unknown parameters, rank $(A) = r \leq m < n$, $p \leq n - r$ and Σ is an unknown dispersion matrix. Under this model suppose we have a testable hypothesis (the meaning and mathematical criterion for testability being discussed in the paper) $H_0 : C(s \times m)\xi(m \times p)M\ (p \times q) = 0\ (s \times q)$ where C and M (to be called the hypothesis matrices) are given such that rank $(C) = s \leq r$ and rank $(M) = q \leq p$. The alternative is $H : C\xi M = \eta(s \times q)$ $(\neq 0)$. The test is that at a level α we accept H_0 if $c_{\max}\ (S^*S^{-1}) \leq c_\alpha$ and reject H_0 otherwise, where S^* and S are matrices given (in the paper) in terms of X, A, C and M, $c_{\max}\ (T)$ denotes the largest root of a matrix with real nonnegative roots, and c_α is a constant depending on α, min (s, q) and $n - r$, which we can pick up from tables now under construction and expected to be published shortly.

MORRIS SKIBINSKY (Michigan State University, Michigan).

447 A Limit Theorem and Bounds for an Optional Stopping Probability.

Let S_j be the standardized jth partial sum of a sequence of bounded independent, identically distributed random variables, K, a positive constant, and let

$$Q(m, n, K) = Pr\{\max_{m \leq i \leq n} S_i \geq K\}$$

It is shown by elementary methods that if $\lim_{m \to \infty} ((n - m)/\sqrt{m}) = 0$, then $\lim_{m \to \infty} Q(m, n, K) = 1 - \phi(k)$, where ϕ is the standard normal c.d.f. Certain steps in the proof are then used to obtain simple bounds for $Q(m, n, K)$ when the sequence of random variables is generated from Bernoulli trials.

EVAN J. WILLIAMS (North Carolina State College, North Carolina, U.S.A.).

448 On Statistics Independent of a Sufficient Statistic.

It is shown that if, for a sample drawn from a population of values of x with distribution depending on a parameter θ, the statistic z is sufficient for θ, and g is any statistic whose distribution is independent of θ, then g and z are independently distributed. The method of proof is less sophisticated than that of Basu (Sankhyā, *15*: 377 [1955]).

The result has application to the normal distribution: the mean of a sample is distributed independently of any location-free statistic; and to the gamma distribution: the mean of a sample is distributed independently of any scale-free statistic. These well known results follow

since the sample mean is a sufficient statistic, in the former case for the location parameter, in the latter case for the scale parameter.

The limitations of the general result lie in the difficulty of deriving statistics independent of parameters other than location and scale parameters.

The connexion of the theorem with estimation theory is discussed.

MARVIN ZELEN (Statistical Engineering Laboratory, Washington 25, D. C.).

449 The Use of Incomplete Block Designs for Asymmetrical Factorial Arrangements.

Let A_s ($s = 1, 2, \cdots, m$) denote the sth factor in a m-factor factorial experiment such that A_s has m_s levels. Let $i = (i_1, i_2, \cdots, i_m)$ represent a particular experimental combination of the m-factors and let the mathematical model underlying the measurements be

$$y_{ij} = \mu + \sum_{s=1}^{m} (a_s)_{i_s} + \sum_{t=2}^{m} \sum_{s=1}^{t} (a_{st})_{i_s i_t} + \cdots + (a_{12\cdots m})_{i_1 i_2 \cdots i_m} + b_j + \epsilon_{ij}$$

where $(a_s)_{i_s}$, $(a_{st})_{i_s i_t}$, $\cdots$, $(a_{12\cdots m})_{i_1 i_2 \cdots i_m}$ represent the various main effects and interactions, b_j represents the block effect, and the ϵ_{ij} are NID $(0, \sigma^2)$. Algorithms are given for using the balanced incomplete and the group divisible designs for asymmetrical factorial arrangements. Let $M(s)$ be the square matrix (of dimension m_s) $M(s) = m_s I - J$ where J is a matrix having all elements unity, and define the direct product of p such matrices by $M(1, 2, \cdots, p) = [M(1) \times M(2) \times \cdots \times M(p)]$ $(p \leq m)$. Then the variance-covariance matrix of a p-factor interaction for the G.D. case can be written as $M(1, 2, \cdots, p)\, \sigma^2/(E_t r v)$ $t = 1$ or 2. For the $BIBD$, the same expression holds with $E_1 = E_2$. The correlations between the different interactions are all zero and since $M^2(1, 2, \cdots, p) = M(1, 2, \cdots, p) \prod_1^p m_s$, $[E_t r / \prod_1^p m_s] \sum (a_{12\cdots p})^2 i_{12\cdots p}$ follows a $\sigma^2 x^2$ with $\prod_1^p (m_s - 1)$ degrees of freedom under the hypothesis of no p-factor interaction effects.

THE BIOMETRIC SOCIETY

International

International Statistical Institute

The 30th meeting of the International Statistical Institute was held in Stockholm from August 8—14, 1958. Three of the sessions were organized by the Biometric Society, and the following papers were presented: *Problems of Experimentation* (organizer, B. Matérn): G.E.P. Box, Elucidation of basic mechanisms; O. L. Davies, Screening tests in the pharmaceutical industry; D. J. Finney, Statistical problems of plant selection; S. H. Justesen and M. Keuls, Note on nonorthogonal designs. *Statistical Genetics* (organizer, L. L. Cavalli-Sforza): J. A. Böök, Practical applications of the theory of population genetics to man; R. A. Fisher, Polymorphism and natural selection. *Statistics in Medical Research* (organizer, L. Martin): J. Cornfield, J. Steinfeld and S. Greenhouse, Models for medical experiments using isotopically labelled tracers; B. G. Greenberg and A. E. Sarhan, Some applications of order statistics; J. O. Irwin, Experiments on the carcinogenic action of mineral oils; A. W. Kimball, Disease incidence estimation in populations subject to multiple causes of death; D. Schwartz and G. Anguera, Une cause de biais dans certaines enquêtes médicales; la temps de séjour a l'hopital; W. Billewicz, Practical problems in a sequential medical trial.

Regional

Région Belge

La Société Adolphe Quetelet, présidée par Monsieur R. Laurent, a eu le plaisir de recevoir à sa tribune, le 27 mars dernier, Monsieur Raymond van den Driessche, Ingénieur des Eaux et Forêts A. I. Gx. et Chef de Travaux de la Division de Biometrie de l'I.N.E.A.C. (Yangambi, Congo Belge). Cette conférence qui eu lieu à la Fondation Universitaire, portait sur le sujet suivant: "La place dans l'expérimentation agricole des plans en blocs incomplets partiellement équilibrés."

British Region

A meeting held on May 30th, 1957 was devoted to problems of human blood groups, the following papers being presented:—J. A. Fraser Roberts, Blood Groups and Susceptibility to Disease; J. H. Renwick and Sylvia Lawler, Mapping Human Chromosomes; A. E. Mourant, The Nature and Possible Interpretation of Blood Group Population Data.

The summer meeting of the Region was held on July 11th, 1957, when a visit was paid to the Central Veterinary Laboratory, Weybridge.

Région Française

A une réunion tenue le 19 juin, Mm. J. Sutter et I. — N. Toan ont présenté une communication entitulée — Problèmes posées par l'étude de la dispersion des gênes dans les populations humaines occidentales.

Switzerland

The Biometric Seminar for Agronomists held in Zurich (July 22–26) at the Swiss Federal Institute of Technology was attended by 31 participants. Introductory lectures on statistical methods were given by Professor Linder (Genf), Professor Lörtscher, Mr. Abt, Dr. LeRoy (Zurich), and Dr. Auer (Zouz).

The seminar was organised by the Swiss Branch of the Biometric Society. Assistance was received from the Swiss Society of Agronomy Engineers.

At the Annual Meeting of the Swiss Branch, held on July 27 in Zurich, the following papers were read: A. Linder (Genf), Die Bedeutung und Berechnung der Streuungskomponenten; Henri L. LeRoy (Zurich), Populations-genetische Betrachtungen zur Erbwertschätzung.

WNAR

The annual meeting was held at Stanford University, August 27 and 28 jointly with the American Institute of Biological Sciences and the AAAS (Pacific Division). The invited address, given by Dr. W. L. Smith of Cambridge, was entitled "A random walk problem arising from the study of the flour beetle."

The program included four sessions: *Problems in Discriminatory Analysis* (presided over by S. W. Nash): papers by Rosedith Sitgreaves, A. C. Walker and W. F. Royce; *Epidemiology* (presided over by W. J. Dixon): papers by M. F. Laughlin, H. B. Messinger, M. S. Ahmed and W. F. Taylor; *Contributed Papers* (presided over by D. G. Chapman): papers by J. Gurland and D. G. Chapman; *Genetics* (presided over by F. T. Schultz): papers by B. O. Berg, W. Becker, H. Abplanalp, R. W. Allard, E. Novitski, Dorothy Lowry and A. C. Walker.

Membership

The paid-up membership of the Society at the end of 1956 was 1303, divided between regions as follows:

Australasian	65	India	14
Belgian	60	Italian	50
Brazilian	68	Japan	47
British	162	Netherlands	33
Denmark	14	Sweden	14
ENAR	473	Switzerland	22
French	67	WNAR	96
German	69	At large	49

Members

The following notifications of changes of address and of location of new members were received during May-July, 1957.

New Addresses

Mr. John C. Bain, 32 South Munn, East Orange, N. J., U.S.A.

Dr. Samuel H. Brooks, 1260 Decisadero, Pacific Grove, California, U.S.A.

Mr. Victor Chew, Department of Experimental Statistics, North Carolina State College, Box 5457, Raleigh, North Carolina, U.S.A.

Dr. Ralph E. Comstock, Department of Animal Husbandry, Institute of Agriculture, University of Minnesota, St. Paul 1, Minnesota, U.S.A.

Dr. William S. Connor, Statistical Engineering Lab., National Bureau of Standards, Washington 25, D. C., U.S.A.

Mr. Ellsworth B. Cook, Naval Med. Field Res. Lab., Camp Lejeune, N. C., U.S.A.

Dr. L. Otis Emik, Air Pollution Med. Prog., Public Health Service, DHEW, Washington 25, D. C., U.S.A.

Dr. Gerald Friedman, 65 Park Terrace East, Apartment C-66, Manhattan, New York 34, N. Y., U.S.A.

Mr. Roger Heimlich, 222 Saf., Columbus, Indiana, U.S.A.

Dr. Henry Klaunberg, National Press Building, Washington 4, D. C., U.S.A.

Dr. Eschscholtzia L. Lucia, c/o Mrs. H. B. Stevenson II, 1358 Weller Way, Sacramento 18, California, U.S.A.

Judson McGuire Jr., Apartada 654, Camaguez, Cuba

Dr. A. M. Mood, General Analysis Corp., 11753 Wilshire Blvd., West Los Angeles, California, U.S.A.

Mr. Jack Moshman, Council for Economic and Industry Research, Inc., 1200 Jefferson Davis Highway, Arlington 2, Virginia, U.S.A.

Mr. Sydney I. Neuwirth, 71–08 Parsons Boulevard, Flushing, New York, U.S.A.
Dr. Bernard Ostle, c/o W. E. Boyes, Dept. 1290, Sandian Corporation, Sandia Base, Albuquerque, New Mexico, U.S.A.
Dr. Earl R. Rich, Dept. of Zoology, University of Miami, Coral Gables 46, Florida, U.S.A.
Miss Jean Roberts, 8103 Kentburg Drive, Bethesda 14, Maryland, U.S.A.
Dr. David W. Robertson, Dept. of Agronomy, Colorado State University, Fort Collins, Colorado, U.S.A.
Mr. O. K. Sagen, 133 Carroll Street, SE, Washington 3, D. C., U.S.A.
Morton D. Schweitzer, Ph.D., 600 West 168 St., New York 32, N. Y., U.S.A.
M. C. K. Tweedie, Department of Mathematics, The University, Manchester 13, Great Britain.
Prof. David F. Votaw, J. R., Sheffield Hall, Yale University, New Haven, Connecticut, U.S.A.
Dr. John E. Walsh, 5019 Donna Avenue, Tarzana, California, U.S.A.

New Members

Belgian

H. A. L. J. Amand, I.N.E.A.C., Yangambi, Belgian Congo
P. Dineur, I.N.E.A.C., Gandajika, Kasai, Belgian Congo
Dr. O. Lazer, 75a, Bd. de Cabbeck, Tirlemont, Belgium
M. J. W. Luttgens, Yangambi, B.P. 37, Belgian Congo
T. J. Marymen, I.N.E.A.C., Gandajika, Kasai, Belgian Congo

British

Dr. Allen Birnbaum, Mathematical Department, Imperial College, Exhibition Road, London, S.W. 7, England
Mr. R. M. Cormack, Department of Statistics, Marischal College, Aberdeen, Scotland
Mr. R. C. Elston, 161 Queen's Drive, London, N. 4, England
Dr. Authur L. Jolly, Imperial College of Tropical Agriculture, St. Augustine, Trinidad, B.W.I.
Mr. C. D. Kemp, The Grassland Research Station, Hurley, Near Maidenhead, Berks, England
Mr. W. L. B. Nixon, Genetics Unit, Institute of Psychiatry, Denmark Hill, London, S.E. 5, England

ENAR

Mr. Ross Willard Adams, 401 Natural History Building, University of Illinois, (Dept. of Physiology), Urbana, Illinois, U.S.A.

Dr. Chester Alexander, Westminster College, Fulton, Missouri, U.S.A.

Mr. Earl L. Atwood, 7117 Woodland Avenue, Takoma Park 12, Maryland, U.S.A.

Dr. Donald W. Bailey, R. B. Jackson Memorial Laboratory, Bar Harbor, Maine, U.S.A.

Dr. Harold J. Bornhold, 27 East Eleventh Street, New York 3, N.Y., U.S.A.

Dr. Hugh Bishop Cannon, Horticulture Division, Central Experimental Farm, Ottawa, Ontario, Canada

Dr. C. K. Chai, Jackson Laboratory, Bar Harbor, Maine, U.S.A.

Dr. Ernest L. Corley, Dept. of Dairy Husbandry, University of Wisconsin, Madison, Wisconsin, U.S.A.

Mr. Richard J. Daum, Department of Entomology, Cornell University, Ithaca, New York, U.S.A.

Mr. Howard H. Engelbrecht, U. S. Weather Bureau, Friendship International Airport, Baltimore, Maryland, U.S.A.

Mr. Peter W. Frank, Department of Zoology, University of Missouri, Columbia, Missouri, U.S.A.

Mr. John J. Gart, Department of Statistics, Virginia Polytechnic Institute, Blacksburg, Virginia, U.S.A.

Prof. Dr. James A. Hagans, Experimental Therapeudic Unit, Dept. of Medicine, University of Oklahoma School of Medicine, Oklahoma City, Oklahoma, U.S.A.

Mr. Jimmy H. K. Kan, Department of Poultry Science, Texas A and M College, College Station, Texas, U.S.A.

Miss Katherine Baker Laughton, 174 Pearl Street, Burlington, Vermont, U.S.A.

Harry H. Shorey, Dept. of Entomology, Cornell University, Ithaca, N.Y., U.S.A.

Dr. J. N. Spuhler, Institute of Human Biology, University of Michigan, Ann Arbor, Michigan, U.S.A.

Dr. N. T. Werthessen, S.W. Foundation for Res. and Education, San Antonio, Texas, U.S.A.

Miss Joan Wax, Research Laboratories, Parke, Davis and Co., Detroit 32, Michigan, U.S.A.

WNAR

Robert W. Allard, 223 Hunt Hall, University of California, Davis, California, U.S.A.

German

Hilmar Grimm, Jena/Thur DDR, Prof.—Ibrahim—Str. 16, Germany
Vladimir Maly, Karlovo nam 32, Prague, CSR
Prof. Dr. Heinz Rethmann, Kiel/Holstein, Esmarchstr. 42, Germany
Erwin Roth, (13a) Markbreit/Main, Enheimer Strasse 481, Germany
Prof. Dr. Wilhelm Rudorf, Max-Planck Institut fur Zuchtungsforchung, Voldagsen uber Elze/Hann, Britische Zone, Germany
Dr. August Schaaf, Schillerstr. 9, Berlin-Eichwalde (DDR), Germany
Karl Schrimpf, Institut fur Pflanzenvau und Pflanzenzuchtung, Stuttgart, Hohenheim, Germany

Netherlands

D. K. deJongh, Jacob van Ruysdaellaan 25, Heemstede, the Netherlands

Indian

Mr. K. C. Agrawal, Senior Statistical Assistant, Government Horticultural Research Station, Saharanpur, India
Mr. Y. N. Bali, Statistical Officer, Cane Commissioners Office, U.P., Lucknow, India
Mr. Rughuvir Giri, Statistician, Directorate of Land Records, Madhya Pradesh, Gwalior, India
Mr. G. P. Kapoor, Deputy Cane Commissioner, U.P. Lucknow, India
Dr. P. S. Lamba, Principal, Rafi Ahmed Kidwai Agricultural Institute, Sehore (Bhopal), M.P., India
Dr. B. Mukerji, Director, Central Drug Research Institute, Chattar Manzil Palace, Lucknow, India
Mr. D. P. Singh, Director, Planning Research and Action Inst., U.P., Kalakankar House, Lucknow, India

Italian

Dr. Ruggero P. N. Ceppellini, Instituto Sieroterapico Milanese, Via Darwin 22, Milano, Italy
Dr. J. Gani, Department of Mathematics, University of Western Australia, Nedlands, W. A., Australia
Dr. Giulio Alfredo Maccacaro, Instituto di Igiene, Via Forlanini 1, Pavia, Italy
Dr. U. Prota, Instituto di Patologia Vegetale, Pizza Conte Moriana 8, Sassari, Italy
Dr. Marcello Siniscalco, Inst. Genetics, Univ. of Naples, Messocannone, 8, Naples, Italy

Japanese

Prof. Masahiro Aya, Agricultural Chem. Insp. Station, 844, Suzuki-shinden, Kodaira-machi, Kitatama-gun, Tokyo, Japan

Professor Shiro Chikuni, Naikai Reg. Fisheries Res. Lab., c/o Fisheries Lab. 1328, Ujina-machi, Hiroshima, Japan

Mr. Akio Kudo, Faculty of Science, Kyushyu University, Hakazaki, Fukuoka, Japan

NEWS AND ANNOUNCEMENTS

Members are invited to transmit to their National or Regional Secretary (if members at large to the General Secretary) news of appointments, distinctions, or retirements and announcements of professional interest.

Dr. Bernard Ostle, formally Professor of Mathematics and Director of the Statistical Laboratory at Montana State College, is now with the Reliability Department of Sandia Corporation, Albuquerque, New Mexico. His mailing address is 3120 Cardenas Drive N.E., Albuquerque, New Mexico.

EDUCATIONAL TESTING SERVICE FELLOWSHIPS

The Educational Testing Service is offering for 1958–59 its eleventh series of research fellowships in psychometrics leading to the Ph.D. degree at Princeton University. Open to men who are acceptable to the Graduate School of the University, the two fellowships each carry a stipend of $2,650 a year and are normally renewable. Fellows will be engaged in part-time research in the general area of psychological measurement at the offices of the Educational Testing Service and will, in addition, carry a normal program of studies in the Graduate School.

The closing date for completing applications is January 3, 1958. Information and application blanks may be obtained from: Director of Psychometric Fellowship Program, Educational Testing Service, 20 Nassau Street, Princeton, New Jersey, U.S.A.